# Cloning and molecular characterisation of the zebrafish colourless gene

Pauliny, Angela

*Award date:*
2002

*Awarding institution:*
University of Bath

# Cloning and molecular characterisation of the zebrafish *colourless* gene

Submitted by Angela Pauliny
for the degree of PhD of the University of Bath
2002

## Copyright





ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

This work is dedicated to my family
and especially to my parents who always gave me great support
and let me follow my endeavours.

## Acknowledgements

I would like to thank Dr. Michelle Southard-Smith (NHGRI, NIH, Bethesda, Maryland, USA) for providing mouse *Sox10* probes and the Zooblot protocol and Dr. Evelyn Chiang (Institute of Molecular Biology, Taipei, Taiwan) for sharing the *sox9a* and *sox9b* clones with me prior to their publication. The *dlx2* probe was kindly provided by Dr. Craig Miller (Institute of Neuroscience, University of Oregon, USA) and *crestin* by Dr. Paul Henion (Neurobiotechnology Centre, The Ohio State University, USA). I would also like to thank Dr. James Lister and Dr. David Raible (Department of Biological Structure, University of Washington, Seattle, USA) for sharing their pCSHSP heatshock vector. Many thanks go out to Prof. Laurence Hurst (University of Bath, England) for phylogenetic analysis of *sox10* and to Matt Brady (NHGRI, NIH, Bethesda, Maryland, USA) who carried out the *in vitro* luciferase assays on wild-type and mutant *sox10* constructs.

I would like to express my gratitude to Prof. Stuart Reynolds (University of Bath, England), who introduced me to my future PhD supervisor Dr. Robert Kelsh (University of Bath, England) and arranged funding for my project. Also, I owe most of my technical and experimental knowledge to Dr. Stone Elworthy (University of Bath, England), who demonstrated with most helpful supervision that molecular biology can work sometimes. I would like to thank our fish technicians John James, Kim Denny and Leanne Price for keeping the facility running smoothly, Kirsten Dutton for frequently watching over "my" fish stocks and babies and Dr. Stone Elworthy for sharing his identified $cls^{m618}$ stocks with me. Dr. Bill Bennett (University of Bath, England), our computer guru, certainly deserves special thanks for persistent battling with Microsoft Windows and crashing computers, and critical discussions.

On a more personal note, I am forever grateful to Rocky Kahn for sharing Overleaze and all its luxuries with me, giving me great confidence and making me believe in myself. I would also like to thank Dr. Marika Charalambous (University of Bath, England) for memories of great evenings with live bands in the Bell and live Jazz in The Old Farmhouse, teaching me how to make the world's best Feta dip and taking care of Felix, my budgie, when I went on holidays. I am also immensely grateful to Prof. Rod Scott, Dr. Alex Jeffries, Dr. James Doughty, Sam Boundy, Candy Au, Andrea Dowling, Dr. Nick Waterfield, Valentina Nikolic and Trevelyan Menheniott (all

University of Bath, England) for keeping me sane and cheerful while writing my thesis by providing welcome distractions. I would like to thank everyone in the teaching lab for their help and advice no matter what subject and especially Felicity Veazey for lending me her laptop computer. Also, a special thanks goes out to Martin Seiter, who printed out all the great colour panels on his fantastic colour laser printer.

Finally, I would like to express my deep gratitude to my supervisor, Dr. Robert Kelsh, for all his guidance, support, helpful supervision, inspiration and patience when correcting my thesis, and everyone in lab 0.76 for providing a fun environment.
The project was funded within the means of a Demonstratorship by the University of Bath, England.

# Table of contents

# List of Figures

# List of Tables

# List of Appendices

# Abbreviations

| | |
|---|---|
| 3'CDS | 3'cDNA Synthesis Primer |
| 3GSP | 3' Gene Specific Primer |
| 3'UTR | 3' Untranslated region |
| 5' UTR | 5' Untranslated region |
| 5'CDS | 5'cDNA Synthesis Primer |
| 5GSP | 5' Gene Specific Primer |
| ABF2 | Yeast ARS-binding protein 2 |
| BAC | Bacterial Artificial Chromosome |
| BMP4 | Bone morphogenetic protein 4 |
| BSA | Bovine serum albumin |
| CDHS | Craniofacial Deafness Hand Syndrome |
| cDNA | complementary DNA |
| *cls* | *colourless* |
| cM | centiMorgan |
| CNS | Central nervous system |
| Col2a1 | Collagen type II |
| cR | centiRays |
| dct | Dopachrome tautomerase |
| DDC | Duplication-Degeneration-Complementation model |
| DEPC | diethylpyrocarbonate |
| DMSO | dimethylsulphoxide |
| dNTP | deoxynucleotide triphosphate |
| *Dom* | *Dominant megacolon* |
| dpf | days post fertilisation |
| DRG | Dorsal root ganglion |
| E8 | Embryonic day 8 |
| EDN3 | Endothelin 3 |
| EDNRB | Endothelin receptor B |
| EMC | early migrating crest |

| | |
|---|---|
| EN1 | Engrailed 1 |
| eng | Engrailed |
| ENS | Enteric nervous system |
| ENU | Ethylnitrosourea |
| F1, F2 | Family 1, Family 2 |
| FGF | Fibroblast growth factor |
| FISH | fluorescent *in situ* hybridisation |
| fkd6 | forkhead-related protein 6 |
| GAL4DBD | Gal4 DNA-binding domain |
| GDNF | Glial cell-line derived neurotrophic factor |
| GSP | gene specific primer |
| HH stage | Hamburger & Hamilton stage |
| HMG | High-mobility group |
| hpf | hours post fertilisation |
| HSCR | Hirschsprung disease |
| IPTG | isopropylthiogalactoside |
| LEF-1 | Lymphoid enhancer binding factor-1 |
| LG | Linkage group |
| LMC | late migrating crest |
| LN54 | **L**oeb, **N**IH, **5**000 rad, **4**000rad |
| Long UP | Long Universal Primer |
| *ls* | *lethal spotting* |
| MATA1 | Yeast mating-type protein A1 |
| *mi* | *microphthalmia* |
| MITF | Microphthalmia transcription factor |
| N3GSP | Nested 3' Gene Specific Primer |
| N5GSP | Nested 5' Gene Specific Primer |
| *nac* | *nacre* |
| NBT/BCIP | Nitro Blue Tetrazolium/5-Bromo 4-Chloro 3-Indolyl Phosphate |
| NMR | Nuclear magnetic resonance |
| NUPM | Nested Universal Primer Mix |
| OCA | Oculocutaneous albinism |

| | |
|---|---|
| Oct3/4 | Octamer-binding protein 3/4 |
| PAC | P1-derived Artificial Chromosome |
| Pax3 | Paired box transcription factor 3 |
| PNS | Peripheral nervous system |
| prim | primordium |
| RACE PCR | (Rapid Amplification of cDNA Ends-Polymerase Chain Reaction |
| RAPD | Random Amplified Polymorphic DNA |
| RET | Rearranged during transfection |
| RFLP | Restriction Fragment Length Polymorphism |
| RO | Reverse osmosis |
| rpm | revolutions per minute |
| RT | Reverse Transcriptase |
| RT-PCR | Reverse Transcription-Polymerase Chain Reaction |
| *s* | *piebald* |
| SCG | Superior cervical ganglion |
| SDS | Sodium dodecylsulfate |
| Short UP | Short Universal Primer |
| $s^l$ | *piebald lethal* |
| *Sl* | *Steel* |
| SLF | Steel factor |
| SMART | Switching Mechanism At 5' end of RNA Transcript |
| SRY | sex-determining region of Y chromosome |
| SSC | Saline-sodium citrate |
| SSCP | Single-Strand Conformational Polymorphism |
| SSLP | Simple Sequence-Length Polymorphism |
| STS | sequence-tagged sites |
| TCF-1 | T cell-specific transcription factor-1 |
| TET HMG | Tetrahymena HMG protein |
| Tm | Melting temperature |
| trp1,2 | tyrosinase related protein 1,2 |
| TUNEL | Terminal deoxynucleotidyl transferase-mediated dUTP nick end-labelling |
| UBF | RNA polymerase I transcription factor |

| | |
|---|---|
| UPM | Universal Primer Mix |
| Wnt | Wingless |
| WS1-4 | Waardenburg Syndrome type 1-4 |
| X-gal | 5-Bromo-4-chloro-3-indolyl b-D-galactopyranoside |
| YAC | Yeast Artificial Chromosome |

## Abstract

*colourless* (*cls*) mutants were isolated in two independent ENU mutagenesis screens for recessive zebrafish mutations. All homozygous mutant alleles are embryonic lethal, whereas heterozygotes are phenotypically wild-type. *cls* mutants present a lack of all 3 types of pigment cells, intestinal aganglionosis and a severe reduction in most other neurons and glia of the peripheral nervous system, in combination with an inner ear defect. The homozygous *cls* phenotype is reminiscent of homozygous *Dominant megacolon* (*Dom*) mouse mutants, a model for the human congenital Waardenburg-Shah Syndrome (WS4). Mutations in the transcription factor Sox10, a member of the highly conserved *Sox* protein family characterised by their HMG-type DNA binding domain, have been identified in *Dom* ($Sox10^{Dom}$) mice and in some WS4 patients. Consequently, we proposed that *sox10* was a good candidate gene for the *cls* locus.

To verify this proposal, we cloned a *sox10* homologue by RT-PCR and RACE PCR and confirmed its true identity by sequence alignment, phylogenetic and *in situ* hybridisation analysis. We demonstrated that *sox10* expression was disrupted in *cls* mutants. Furthermore, we showed that the *cls* locus and *sox10* were tightly linked on the distal end of LG 3, and identified mutant lesions in *sox10* in 3 of the *cls* mutant alleles. Ectopic expression of wild-type *sox10* under the control of a heatshock promoter rescued melanophores to a wild-type morphology and migrational behaviour, whereas equivalent constructs expressing mutant proteins failed to do so. Taken together, our results very strongly support the hypothesis that *cls* encodes *sox10*.

Furthermore, we investigated whether the expression patterns of *sox10* and its closest homologues *sox9a* and *sox9b* overlapped and whether either of the Sox9 proteins might show functional redundancy with Sox10, at least in the melanophore lineage.

# Chapter 1 – Introduction

## *Zebrafish as a model organism*

The zebrafish, *Danio rerio* (formerly known as *Brachydanio rerio*), belongs to the Super-Order Teleostei (bony fish) and the Family Cyprinidae. It is native to the rivers of India and Pakistan where it is found in rather shallow waters of 20-28°C. Zebrafish live socially in schools. The adult fish is about 4-5cm long, males tend to be of a slimmer build than females and richer in colour with a slight yellowish or reddish undertone especially on the anal fin (Riehl and Baensch, 1996).

Adults usually spawn in early morning where females lay up to 300 transparent eggs approximately 0.7mm in diameter. Larvae hatch at 2-3 days after fertilisation and start to swim and feed freely after 5 days. Maturity is reached after approximately 3 months depending on food supply and water temperature.

Zebrafish are widely used for toxicological studies in industrial laboratories. Only recently have its advantages as a model organism in developmental and genetic studies been fully appreciated. Zebrafish breed easily in captivity and can be kept fairly cheaply at high densities. The eggs are fertilised externally and their embryos are transparent, which enables cell biological studies in the developing embryo such as single cell labelling with vital dyes. Likewise, analysis of *in situ* mRNA hybridisations in whole-mount embryos is possible.

The advantages for genetic studies are large family size, the short life cycle of approximately 90 days and ease with which recessive mutations can be created by placing males into a mutagen solution containing the point mutation inducing substance ethylnitrosourea (ENU).

In teleosts, a fairly recent, in evolutionary terms, genome duplication was proposed to have taken place more than 100 million years ago (Amores et al., 1998). Duplicated genes might have been lost during evolution unless they evolved to develop new functions or acquired new regional or temporal specificity. Thus, for every gene in mouse one has to consider the possibility of two homologues in zebrafish that have split the gene's original function between themselves. This might be a disadvantage to this model organism, since potentially two such paralogues have to be isolated and analysed.

However, the duplication, degeneration, complementation (DDC) model predicts that duplicates may become preserved by degenerative loss of complementary gene subfunctions from the duplicate copy (Force et al., 1999). Hence, a complex mammalian gene function might be shared complementarily by two zebrafish homologues. For example, in mammalian mutants, a late gene function might be obscured by defects or premature death of the embryo caused by the gene's essential early function. In zebrafish however, this might be overcome by the separation of these two functions (Postlethwait et al., 1999).

Zebrafish are a particularly suitable model organism to study neural crest development. Apart from the general advantages mentioned above like easy access and transparency of embryos, the neural crest cells in zebrafish are significantly larger and fewer in number compared to neural crest cells in avian or mammalian embryos (Raible et al., 1992) facilitating cell biological manipulations.

## *Large-scale mutagenesis screens in zebrafish for recessive mutants*

Systematic mutagenesis screens were first carried out in invertebrates such as *Drosophila* (Nuesslein-Volhard et al., 1984; Wieschaus et al., 1984) and *C. elegans* (Kemphues et al., 1988). However, large-scale screens in vertebrate organisms such as mouse are difficult due to high costs involved in keeping and breeding such large numbers of animals.

In contrast, zebrafish, with its short life cycle, large family size, small space requirements and transparent externally fertilised embryos proves to be a very suitable vertebrate model system in which a large-scale saturation screen can be attempted. Mullins has described in detail the most efficient and reliable method to conduct such a mutagenesis screen (Mullins and Nuesslein-Volhard, 1993) using ethylnitrosourea (ENU) as the mutagen of choice (Mullins et al., 1994).

To date, two major mutagenesis screens have been reported. One was carried out in the Massachusetts General Hospital by W. Driever and his colleagues (Driever et al., 1996). They screened 2383 mutations affecting embryonic and early larval development, resulting in identification of 220 genetic loci. In parallel in Tuebingen, Germany, C. Nuesslein-Volhard and her colleagues identified a total of 4264 mutants of which 1163 were characterised in more detail. By complementation crosses, 894

mutants were assigned to 372 genes which were important in various developmental processes (Haffter et al., 1996a). Mutant phenotypes identified in the screen were organised into many phenotypic groups such as mutations affecting early development, formation of body axes, development of organs, pigment cells and mutations with an adult phenotype (Haffter et al., 1996a). Most of the mutations isolated were embryonic lethal. Since ENU is known to cause mostly point mutations or small deletions, embryonic lethality indicated that the function of these genes was indispensable in development and thus no other gene was able to fully compensate for the loss. However, with an average allele frequency of 2.4 it was statistically possible to obtain mutations located in various regions of the same gene. Thus, for some loci mutants with weaker phenotypes were identified; for others, adult viable phenotypes were created making the study of the gene's role in development a lot easier. The 79 isolated adult viable and 19 semiviable mutations were described in more detail in (Haffter et al., 1996b).

## *Neural crest development*

Although the neural crest is derived from the ectoderm it has often been referred to as the fourth germ layer. A discrete set of originally multipotent cells eventually become specified and differentiate into an incredibly diverse range of derivatives. The neural crest was first described in avian embryos by His (His, 1868 cited in Le Douarin and Kalcheim, 1999) as a strip of cells located between the neural tube and the overlying dorsal ectoderm, which he called "Zwischenstrang". Since then, techniques to identify, visualise and trace single neural crest cells have led to a better understanding of the underlying processes of neural crest development not only in avians, but also in amphibians, fish and mammals. Even so, a lot of progress has yet to be made to resolve details of the mechanisms controlling neural crest development.

### The formation of the neural crest

The neural crest delaminates from the dorsal neural tube in a rostrocaudal manner. Neurulation in avians, amphibians and mammals involves the induction of the neural plate from the ectoderm, followed by an elevation of the lateral margins of the neural plate called neural folds, which finally fuse at the dorsal midline to give rise to the

neural tube. In contrast, the neural tube in teleosts like the zebrafish and in caudal regions of avians is formed by a process called secondary neurulation, whereby a ventral thickening of the ectoderm can be observed. This structure, called a neural keel, then cavitates to give rise to the neural tube (Lamers et al., 1981). The neural crest cells delaminate from the lateral edges of the neural keel and converge towards the dorsal midline (Raible et al., 1992; Kimmel et al., 1995; Thisse et al., 1995; Baker and Bronner-Fraser, 1997b; Le Douarin and Kalcheim, 1999).

The molecular mechanisms underlying neural induction and the subsequent formation of the neural crest are still not fully understood. During neurulation the ectoderm gets divided into non-neural ectoderm and neural plate, which give rise to the epidermis and the neural tube, respectively. This process is thought to be regulated by a gradient of bone morphogenetic protein-4 (BMP-4; an epidermalising factor) originating from ectodermal cells, which is generated by a gradient of a BMP inhibitor (neuralising factor) such as noggin and chordin (Smith and Harland, 1992; Piccolo et al., 1996) secreted by the future axial mesoderm (organizer region). Dorsal ectoderm, which by default adopts a neural fate (reviewed in Weinstein and Hemmati-Brivanlou, 1997), would be epidermalised by BMP-4 if factors like noggin and chordin did not prevent BMP-4 from binding to its receptor (Piccolo et al., 1996; Zimmerman et al., 1996). Hence, epidermis is only formed furthest from the organiser (future axial mesoderm and notochord), whereas neural tube is formed adjacent to this source of noggin and chordin. The neural crest forms at the interface between neural plate and non-neural ectoderm.

There are several models of neural crest induction (reviewed by Baker and Bronner-Fraser, 1997b).

Albers proposed (Albers, 1987) that a change in ectodermal competence to a signal from the organiser could give rise to neural plate, the neural crest and finally epidermis. Raven and Kloos were the first scientists to suggest that neural crest induction might involve signals from the archenteron roof (paraxial mesoderm) directly underlying the prospective neural crest region but, most importantly, that this induction was independent from the induction of the neural plate (Raven and Kloos, 1945). They could show in grafting experiments that medial archenteron roof was able to induce both neural plate and neural crest markers from overlying competent ectoderm, whereas lateral archenteron roof induced neural crest markers only. Marchant and colleagues further investigated these proposed inducing abilities of mesoderm. They conclude the

existence of a neural crest inducer originating from the dorsal (axial) mesoderm creating a gradient along the dorsolateral (paraxial) mesoderm. *In vitro,* dorsal mesoderm was able to induce neural plate markers like *Xsox2* proximally and neural crest markers like *Xslug* at a distance. If dorsal mesoderm was replaced by dorsolateral mesoderm in those conjugates, neural crest markers were expressed next to the inducing tissue, whereas ablation of dorsolateral mesoderm in early gastrula embryos resulted in loss of the neural crest population (Marchant et al., 1998).

A third model suggests that interactions between the neural plate and adjacent non-neural ectoderm (future epidermis) induce the neural crest. This has been indicated by grafting experiments in amphibians (Rollhauser-ter Horst, 1980; Moury and Jacobson, 1990) and avians (Selleck and Bronner-Fraser, 1995). Juxtaposed neural plate and ventral epidermis were found to be sufficient to generate neural crest derivatives in axolotl embryos. Although both tissues contributed to the generation of neural crest cells, most neural crest formed from the neural plate gave rise to melanocytes, whereas epidermal cells formed neurons of the dorsal root ganglion (Moury and Jacobson, 1990).

Can any of these models fully explain the induction of neural crest? BMP-4 is sufficient to repress neural and induce epidermal fates (Wilson and Hemmati-Brivanlou, 1995). Thus, a role in dorsoventral patterning of the ectoderm was suggested by which the neural plate and the non-neural ectoderm could be specified. This is achieved by a delicate balance of BMP and BMP antagonist levels such as chordin and the neural crest could be induced by intermediate levels of BMP-4. This balance can be manipulated by overexpression studies. If BMP-4 is overexpressed in early *Xenopus* embryos by mRNA injection, the neural plate area is reduced, but without affecting actual neural crest formation lining the border (LaBonne and Bronner-Fraser, 1998). Prospective neural crest cells can be identified by their expression of the zinc finger transcription factor *XSlug* (Mayor et al., 1995; LaBonne and Bronner-Fraser, 1998). Thus, BMP-4 seems to be important in defining the location of the neural plate borders. In contrast, overexpression of chordin expands the region of *Xslug* expression laterally at the expense of non-neural ectoderm. However, at no concentration of chordin, and thus of BMP-4, could wild-type levels of *XSlug* expression be observed (LaBonne and Bronner-Fraser, 1998). Furthermore, LaBonne and her colleagues never observed melanocytes being formed in ectodermal explants overexpressing chordin. Taken together, this suggests that BMP might provide an initial weak neural crest specification

at the borders of the lateral neural plate, but additional signals derived from either the adjacent non-neural ectoderm and/or underlying mesoderm may then be required to enhance and maintain neural crest induction at wild-type levels (LaBonne and Bronner-Fraser, 1998).

Members of the Wnt and FGF family have been found to alter the anteroposterior characteristics of noggin-induced neural plate tissue (McGrew et al., 1997). Recently, it has also been shown that there is a direct requirement for Wnt signalling in neural crest formation. In contrast to chordin alone, conjugates of chordin-expressing and XWnt-8 or eFGF-expressing explants were able to induce wild-type expression levels of *Xslug*. Furthermore, a significant number of melanocytes were formed in this chordin-induced neural tissue (LaBonne and Bronner-Fraser, 1998). Since XWnt-8 is a downstream target of eFGF, it was tested whether the induction of neural crest was achieved by eFGF directly and through XWnt-8 signalling. When a dominant negative mutant of XWnt-8 was co-expressed, the ability of eFGF to induce *Xslug* expression in chordin-expressing explants was drastically reduced. This suggested an indirect role for eFGF in neural crest formation, whereas XWnt-8 was shown to have a direct effect on neural crest induction *in vitro* and *in vivo* (LaBonne and Bronner-Fraser, 1998).

Could the role of XWnt-8 in neural crest and somite formation be connected? Paraxial mesoderm (prospective somites) is a potent neural crest inducing tissue. Since a dnWnt has been shown to disrupt somite formation, it is possible that the mechanism by which XWnt-8 acts on neural crest formation involves disrupting mesoderm, which in turn fails to induce neural crest. This hypothesis could be disproved by LaBonne and Bronner-Fraser. They demonstrated that neural crest failed to form even when dnWnt was exclusively expressed in the ectoderm and that the requirement for XWnt-8 in neural crest induction was thus independent of any role it might have in mesoderm (LaBonne and Bronner-Fraser, 1998).

A similar role for a member of the Wnt family in patterning the mediolateral polarity (dorsoventral axis after formation of the neural tube) has been shown for XWnt7B in *Xenopus* (Chang and Hemmati-Brivanlou, 1998). XWnt7B was able to induce neural crest when co-injected with noggin into ectodermal explants. Presumptive neural crest was visualised by staining for *Xslug* and *Xtwist*. This induction still took place after neuralising explants by dissociation and thus even in absence of epidermis. In

overexpression studies *in vivo*, XWnt7B also expanded the expression domain of the neural crest marker *Xtwist*.

In summary, the current model suggests a two-signal induction process leading to neural crest formation. Firstly, BMP levels have to be decreased sufficiently by inhibiting the interaction between BMP and its receptor by neuralising factors like chordin or noggin. The induced neural tissue is of anterior character and can then be caudalised by factors such as FGF.

This decreased level of BMP signalling also specifies the lateral border regions of the neural plate and renders them competent to respond to lateralising (dorsalising) Wnt signals that enhance and maintain neural crest induction. Signals involved in the patterning of the mediolateral character of the open neural plate (future dorsoventral axis of the neural tube) are thought to be derived from the notochord (e.g. hedgehog promoting ventral neural fates like floorplate and motor neurons) and from the non-neural ectoderm (e.g. Xwnt7B mediating dorsal cell fates like neural crest and roofplate).

Overexpression of *Slug* alone appears insufficient to induce neural crest at wild-type levels. However, once *Slug* is induced by BMP/Wnt like signals it seems to regulate its own expression. Thus, Slug appears to have a role in neural crest fate maintenance, additionally to a later role in neural crest migration.

This model is broadly consistent with results of studies in zebrafish. The zebrafish mutant *swirl* was shown to encode *bmp2b* and to be necessary for neural crest formation (Nguyen et al., 1998). *swirl* mutants have a severely dorsalised phenotype, whereas *chordino* mutants, a *chordin* homologue, are ventralised. This is consistent with the interplay of BMP-4 and chordin in dorsoventral patterning of the frog embryo. Aside from ventral cell-types, laterally-derived neural crest progenitors, labelled with *fkd6*, are severely reduced or absent in *swirl* mutants. However, this region is enlarged in *somitabun* (*smad5*) and to a lesser degree in *snailhouse* (*bmp7*) mutants, two other genes of the BMP pathway. The phenotype of these 3 mutants were explained by the activity of a morphogen. Different levels of BMP activity induce differential gene expression along the dorsoventral axis (Nguyen et al., 1998; Barth et al., 1999). This results in the specification of various cell-types, which is also observed in frogs. Thus, to induce neural crest, low levels of BMP (*swirl*) are necessary at the border between the neural plate and the non-neural ectoderm. However, consistent with frogs, it was

proposed that additional factors apart from BMP may be required for neural crest induction (Nguyen et al., 1998).

## Neural crest segregation, specification and migration

The neural crest is induced in bilateral bands of cells at the margin of the neural plate and is labelled by *fkd6* from about 90% epiboly (Odenthal and Nuesslein-Volhard, 1998). Such *fkd6* expressing cells are observed posterior to a region stained with *floating head* (*flh*), which will give rise to the prospective dorsal diencephalon (Masai et al., 1997; Barth et al., 1999).

Neural crest segregation, like induction, occurs in a rostrocaudal sequence. At the anteroposterior level of somite 8 the neural keel starts to segregate from the overlying neuroepithelium at 15hpf and the neural tube has formed from the neural keel by cavitation at 17hpf. In contrast to the non-neural ectoderm, the basal lamina covering the neural tube is discontinuous at this stage and might thus enable the segregation of neural crest cells from the dorsal neural tube. This process is finished by 19hpf and neural crest cells are closely associated with the now continuous basal lamina of the neural tube (Raible et al., 1992). Neural crest cells in zebrafish are approximately twice the size of their avian equivalents and less numerous. In fact, zebrafish on average contain 3-5 times fewer premigratory neural crest cells per segment than avians (Raible et al., 1992).

### *Migration of neural crest cells*

Migration of cranial and trunk neural crest cells show differences in the timing and choice of pathway. In avian embryos, cephalic neural crest cells are located anterior to somite 5 (Baroffio et al., 1991). The first cranial neural crest cells in birds start to migrate as a sheet from above the mesencephalon at the 6 somite stage and follow a subectodermal pathway (Baker et al., 1997a). At the hindbrain level, neural crest cell migration progresses in distinct bands corresponding to branchial arches (Le Douarin and Kalcheim, 1999).

Similarly, cranial neural crest cells in zebrafish just caudal to the developing eye, start to migrate subectodermally at 15-16hpf. Cells from each axial level contribute to a specific pharyngeal segment (Schilling and Kimmel, 1994) and these segment restrictions appear to be generated sequentially in a rostrocaudal manner. Thus, all cells

labelled as early as 12hpf in the rostralmost regions exclusively contribute to the mandibular or hyoid arches, but cells labelled more caudally at the same time migrate into 3 adjacent segments (Schilling and Kimmel, 1994).

In avians, the neural crest population located above or caudal to somite 5 is referred to as the trunk neural crest (Sieber-Blum, 1990). Trunk neural crest cells in vertebrates follow two conserved migrational pathways, the medial (ventromedial) and the lateral (dorsolateral) pathway.

Neural crest cells first enter the medial pathway at approximately 16-17hpf in zebrafish. The onset of neural crest migration in the trunk is linked to the differentiation of the somites. Premigratory neural crest cells are not in contact with the dorsal somite surface, but begin to probe the latter with long pseudopodia (Raible et al., 1992). As somites differentiate they elongate along the dorsoventral axis bringing the dorsal somite surface in contact with neural crest cells. This elevation is concomitant with the start of neural crest migration ventrally between the neural tube and the somite at 18hpf at the level of somite 8 (Raible et al., 1992). In contrast to birds and mammals, cells enter the pathway at any axial level along the somite. By the time they reach the ventral edge of the neural tube they all seem to have converged towards the middle of the somite (Raible et al., 1992). Avian and mammalian neural crest cells however are restricted to the rostral part of the somite (Rickmann et al., 1985; Bronner-Fraser, 1986). Approximately 4 hours after cells have started migration on the medial pathway in zebrafish, some neural crest cells additionally enter the lateral pathway. They disperse from the dorsalmost part of the neural tube and migrate dorsolaterally, then ventrally between the somite and the overlaying epidermis (Raible et al., 1992). Neural crest cells in zebrafish do not seem to be restricted to a region along the somite when they enter the lateral pathway, similarly to observations on the medial pathway. Interestingly, neural crest cells in mouse seemed to enter both pathways simultaneously (Serbedzija et al., 1990), whereas in avians, a 24 hour delay has been reported between the onset of the two pathways (Erickson et al., 1992).

### *Derivatives of the cranial and trunk neural crest*

Cell lineages derived from cranial and trunk neural crest cells are broadly overlapping. However, cartilage precursors are only formed from the cranial neural crest, whereas sympathetic neurons for example, are only derived from the trunk neural crest cell population.

The cell-types derived from the cranial neural crest have been extensively studied by Schilling and Kimmel. By labelling single neural crest cells with a fluorescent dye and recording their fate they were able to construct a fate map of cranial neural crest located in the region between the eye and the first somite (Schilling and Kimmel, 1994). It was shown that neural crest in the mesencephalic and rostral rhombencephalic region consisted of 6 tiers of cells. Whereas the anterioposterior position of neural crest cells determined which pharyngeal segment they would contribute to, the cell-type could be predicted from the mediolateral position. The most lateral tier of cells exclusively gave rise to neurons in cranial ganglia. More medial tiers 2-5 gave rise to Schwann or satellite glial cells and pigment cells, whereas the most dorsally located neural crest cells in tier 4-6 differentiated into cartilage and connective tissues forming head structures (ectomesenchymal fates; Schilling and Kimmel, 1994). This fate map established for the cranial region in zebrafish seemed to roughly correlate to the avian fate map of that region (Noden, 1987). However, a similar study carried out by Dorsky, revealed some discrepancies with Schilling's fate map (Dorsky et al., 1998). When Dorsky labelled cranial neural crest cells, neurons were derived exclusively from a lateral position in agreement with Schilling's study. However, pigment cells mostly arose from medially located neural crest cells and glia from all positions, although slightly fewer from medial positions. Most strikingly, cartilage precursors were labelled equally in all mediolateral positions (Dorsky et al., 1998).

The discrepancies between the two studies might be explained, if different cell layers had been labelled. Schilling states that he only labelled the most peripheral 20% of neural crest cells, whereas Dorsky claims to have labelled representatives of all layers. 88% of all labelled clones seem to be fate restricted prior to the onset of migration and gave rise to a single cell-type. This is very different to data reported for the cranial neural crest population in avians (Baker et al., 1997a). Using isochronic isotopic transplantation of cranial neural crest cells between quail and chicken embryos they showed that all neural crest precursors could give rise to all derivatives regardless of when they started their migration. Thus, in contrast to results reported in zebrafish (Raible and Eisen, 1994; Schilling and Kimmel, 1994; Raible and Eisen, 1996), she concluded that neither early nor late migrating cranial neural crest populations were lineage restricted. Furthermore, early neural crest cells contributed equally to dorsal and ventral fates, whereas derivatives formed by late neural crest cells were confined to dorsal regions and gave rise to fewer ectomesenchymal fates. However, these

transplants in avians consisted of strips of neural crest cells, 10-15 cells wide and spanning the length of the mesencephalon. These grafts likely contained cells of various fate-restrictions and thus might have appeared to be multipotent as a group.

In avian and mammalian embryos, neural crest cells in the trunk give rise entirely to non-ectomesenchymal derivatives such as neurons of the sensory and sympathetic ganglions, Schwann cells and pigment cells. Enteric neurons and glia are derived from three locations in avians. The neural crest above somites 1-7, the vagal region, forms the majority of the enteric nervous system colonising the gut in a rostrocaudal sequence. Neural crest above the lumbosacral spinal cord contributes mostly to enteric ganglia in the colorectum whereas some neural crest originating from the anterior trunk region colonises the esophagus and anterior part of the stomach (Durbec et al., 1996; Burns and Le Douarin, 1998). In zebrafish, trunk neural crest also forms fin mesenchyme, an ectomesenchymal fate (Smith et al., 1994).

Most neural crest cells migrating on the medial pathway in zebrafish are fate-restricted. They give rise to single cell phenotypes, whereas only 20% produce multiple derivative clones (Raible and Eisen, 1994). Labelling single neural crest cells at different time points revealed that the number of neural crest precursors producing multiple phenotypes decreased with time as those precursor cells undergo progressive fate restriction. Sensory and sympathetic neurons are only derived from neural crest cells that start migrating before 18hpf, whereas time of migration is less important for all other derivatives. Raible and his colleagues could also demonstrate that at least some neural crest cells were specified before they reached their final destination. Some neural crest cells, like precursors for dorsal root ganglia (DRG) exhibited identifiable and reproducible migration patterns; others started to synthesize melanin (melanophores).

Neural crest progeny migrating on the lateral pathway consists of type-restricted precursors only and exclusively gives rise to clones of either all melanin-positive (melanophores) or all melanin-negative (xanthophores and iridophores) (Raible and Eisen, 1994).

### *How do neural crest cells become specified?*

There are two models of neural crest development that attempt to explain how this wide variety of different neural crest derivatives is formed. The first one assumes that neural crest cells are predetermined to generate specific cell-types. In the second model all neural crest cells are equivalent initially and later become fate restricted to give rise to

various cell-types by localised environmental signals they encounter during their migration (Raible and Eisen, 1995). Published results seem to indicate that a combination of both, intrinsic differences between cells and environmental cues might be involved in neural crest specification.

Some support for the first model is observed by single cell labelling of premigratory neural crest cells above the zebrafish trunk with lysinated rhodamine dextran. Most trunk neural crest cells on the medial pathway generated derivatives of a single phenotype (Raible and Eisen, 1994). Furthermore, all labelled neural crest cells on the lateral pathway produced clones of pigment cells that were either all melanin-positive or all melanin-negative (Raible and Eisen, 1994). Similarly, 88% of rhodamine dextran labelled premigratory cranial neural crest cells generated progeny of single cell-types (Schilling and Kimmel, 1994). Although results in these studies clearly demonstrate that cells give rise to only a single neural crest fate, they do not show the potential of these cells. It was not investigated, whether the same neural crest cells could still change their fate, if they were placed in a different environment, for example. In fact, there are indications from studies described below that the specification status of neural crest cells *in vivo* is not in agreement with the one predicted by the first model.

The second model proposed a fate restriction for neural crest cells during their development in response to intrinsic or extrinsic cues. Evidence supporting this model was provided by (Raible and Eisen, 1996). They studied cell-type restrictions of trunk neural crest cells migrating on the medial pathway using single cell labelling. They asked, whether early and late migrating populations of neural crest cells had different fate potential. In their observation, 20% of clones migrating early on the medial pathway gave rise to multiple phenotype clones, whereas two hours later, cells migrating at the same axial level only produced single phenotype clones (Raible and Eisen, 1996). To test, whether these cells had undergone progressive fate-restriction and also changed their developmental potential, they used transplantation studies. As an example, they specifically investigated a possible change in neuronal potential. They found that trunk neural crest cells migrating on the medial pathway early (EMC, in segment 7 before 18hpf) gave rise to sensory neurons of the DRG, glia and pigment cells. However, late migrating neural crest (LMC, in segment 7 after 18hpf) had lost their potential to generate sensory neurons under those environmental conditions.
In their experiments, EMC transplanted into older hosts and thus migrating at the same time as endogenous LMC were still able to give rise to DRG neurons, whereas

transplanted into an even later host (21-23.5hpf) they no longer were able to do so. They suggested that a possible change in an environmental signal or a progressive fate-restriction might be responsible for the loss of ability to generate DRG neurons. LMC transplanted into younger hosts did not gain the ability to produce sensory neurons even under the "early" environment, EMC are usually subjected to. However, if such LMC were transplanted into those early hosts after ablation of the EMC population they migrated ventrolaterally to fill in the EMC positions. They were able to give rise to DRG and thus functionally replaced the early neural crest population. Raible and Eisen concluded that LMC and EMC still have the same developmental potential although they exhibit intrinsic conditional biases towards creating specific fates in their normal environment (Raible and Eisen, 1996).

Specification of trunk neural crest derivatives in avians appears to be similar to zebrafish. Henion and Weston (Henion and Weston, 1997) asked whether the composition of fate-restricted and unrestricted precursors changed within a neural crest population over time. They randomly labelled single neural crest cells of a primary neural crest cell culture by injection of a lineage tracer. Cells were grown in a medium permissive for neuronal, glial and pigment cell fates and identified with cell-type specific markers at various time points. They found that the initial neural crest population, which had emerged from the neural tube explant and was labelled during the first 6 hours consisted of both pluripotent (55.5%) and fate-restricted precursors (44.5%). However, 87% of neural crest cells labelled at 30-36 hours after emergence were fate-restricted. This progressive fate-restriction had also been reported for zebrafish trunk neural crest cells on the medial pathway (Raible and Eisen, 1996). They further asked in this study, at what stages different cell lineages separated from each other. The neurogenic lineage seemed to become specified between 13-16 hours and 30-36 hours before the glial and melanogenic lineages and was the first one to start migration. Fate-restricted melanogenic precursors emerged last from the neural tube, but prior to overt differentiation, which was true for all sublineages. Since these fate-restrictions were observed under a presumably uniform cell culture environment, they concluded that intrinsic differences between precursor populations or cell-cell interactions must be responsible for neural crest diversity rather than environmental cues alone. Those distinct subpopulations might then respond to the present environment in different ways promoting separate developmental pathways (Henion and Weston, 1997).

In summary, the studies described above seem to fit a model for neural crest specification that is a combination between the two extreme original models. In this refined model, the initial population of neural crest cells consists of a mixture of fate-restricted and multipotent precursors, which become progressively fate-restricted over time.

How can two adjacent cells with equivalent potential adopt different fates? There are two mechanisms described in the literature. The first involves lateral inhibition, by which cells to be specified complementarily instruct each other; in the second mechanism, cells respond differently to the same signal depending on their distance from it.

Cornell and Eisen investigated how cells with apparently equivalent developmental potential in the lateral neural plate could be specified to form either neural crest cells or Rohon-Beard spinal sensory neurons. They found that cells singled out to be precursors for Rohon-Beard neurons expressed *deltaA*, which bound to the Notch receptor induced in the surrounding cells. Notch-Delta signalling was shown to be necessary to prevent these surrounding cells from adopting the Rohon-Beard fate and allowing them to become neural crest instead (Cornell and Eisen, 2000).

The Wnt signalling pathway on the other hand was shown to be involved in making specific fates from the neural crest. Dorsky presented evidence that suggested a role for Wnt-1 and/or Wnt-3a in specifying pigment cell fates in the cranial region (Dorsky et al., 1998). *Wnt-1* and *Wnt-3a* expressed in the anterior dorsal neural keel seemed to promote pigment fates in adjacent neural crest cells at the expense of neurons and glia. Interestingly, the formation of ectomesenchymal fates was not affected in any way by these Wnt signals further supporting previous data that those fates represented a separate neural crest lineage from the non-ectomesenchymal derivatives such as neurons and pigment cells (Le Douarin and Teillet, 1974).

## *Pigmentation mutants and the colourless mutation*

A great number of mutations isolated in the Tuebingen screen affect neural crest derivatives and especially pigment cells. In total, 285 mutations within 94 distinct genes were found, which affect various aspects of embryonic and larval pigment development such as pigment specification, distribution and survival (Haffter et al., 1996a). The second screen carried out in Boston identified 54 genes (Driever et al., 1996).

Pigmentation mutants from the Tuebingen screen were described in detail by (Kelsh et al., 1996) and mutants were distributed into 7 classes according to their phenotype. Class I describes mutants such as *colourless* (*cls*), with no or strongly reduced numbers of all 3 types of chromatophores, melanophores, xanthophores and iridophores. Mutants in class II show a reduced number of only one chromatophore, e.g. xanthophores or iridophores in *salz* (*sal*) and *shady* (*shd*), respectively, whereas mutants with a reduced number of melanophores, like *sparse* (*spa*), are listed in class III. Class IV contains mutants with altered pigment patterns like *choker* (*cho*), which forms morphologically normal chromatophores in a pattern different to wild-type larvae. *Moonshine* (*mon*) is a representative of Class V mutants, which develop a surplus of chromatophores in ectopic locations. Class VI contains 10 subgroups each describing a reduced chromatophore pigmentation phenotype as can be seen in *albino* (*alb*) embryos, whereas class VII describes mutants with altered chromatophore morphology such as *union jack* (*uni*) (Kelsh et al., 1996).

## The *colourless* mutation in zebrafish

The *colourless* phenotype was initially described by (Kelsh et al., 1996) and (Malicki et al., 1996). *colourless* seemed to have the most dramatic phenotype of all pigment mutants, since it lacked all 3 types of chromatophores and thus suggested an important role during pigment cell development. *cls* first arose spontaneously within the *rose* stock kept in Tuebingen ($cls^{t3}$). It was defined as having no chromatophores apart from a fully pigmented eye and additionally showing an early ear defect. Very few remaining chromatophores were observed in some individuals in normal positions although melanophores, where present, were unusually small and only found on the dorsal stripe. In a more recent study (Kelsh and Eisen, 2000b) it was shown that the pigmentation defect resulted from a decrease of actual melanoblast number, not disrupted melanin synthesis. Fewer melanoblasts were labelled with *dct* in *cls* embryos compared to wild-type siblings at 27hpf and migration of these *dct*-positive cells seemed to be impaired. Furthermore, the number of migrating neural crest cells on the lateral pathway (fated to become all three types of pigment cells; Raible and Eisen, 1994) was drastically reduced in homozygous mutant embryos. However, premigratory neural crest labelled with *fkd6* was formed and maintained in apparently normal numbers in *cls* embryos up to 24hpf.

*cls* embryos also show severe defects in other non-ectomesenchymal fates such as peripheral neurons and peripheral ganglion glia. The number of enteric neurons was severely reduced and enteric glia were either also severely reduced or absent at 7dpf. Sensory neurons of the DRG were reduced anteriorly as early as 2dpf, but absent posteriorly at 7dpf and sympathetic neurons also could not be detected at 9 dpf. Although cranial ganglion neurons appeared normal, satellite glia associated with cranial ganglia and Schwann cells around cranial nerves were markedly reduced in *cls* embryos at 24hpf.

Ectomesenchymal neural crest derivatives such as craniofacial skeletal structures and median fin mesenchyme in head and trunk form normally in *cls* siblings (Kelsh and Eisen, 2000b).

In addition to those defects in neural crest derivatives, *cls* has an early effect on ear development and morphology first detectable around 36hpf (Kelsh et al., 1996; Malicki et al., 1996; Whitfield et al., 1996). By 3dpf the size of the otic capsule is slightly smaller and by 5dpf the tiny ear contains two small otoliths and only one sensory patch (Malicki et al., 1996; Whitfield et al., 1996).

All *cls* mutant alleles show a recessive phenotype. In the Tuebingen screen 4 *colourless* alleles were isolated, called $cls^{ty22f}$, $cls^{te275}$, $cls^{tw2}$ and $cls^{tw11}$.The latter two were derived from the same mutagenised family founder male. In the Boston screen two *cls* alleles were originally classified as mutants with defective ear morphology combined with abnormal pigment pattern. They were named *golas* (*gos*) and the two alleles $gos^{m241}$ and $gos^{m618}$ (Driever et al., 1996; Malicki et al., 1996). $gos^{m241}$ was described to exhibit a weaker phenotype than $gos^{m618}$, but unfortunately the stock seems to have been lost since.

Mutations affecting *cls* and *golas* failed to complement each other and thus are allelic. As a result, $gos^{m618}$ was renamed $cls^{m618}$ (Malicki et al., 1996).

The *cls* gene acts cell-autonomously at least in the pigment cell lineages and homozygous mutants, which will be referred to as *cls* embryos hereafter, die at 10dpf (Kelsh et al., 1996; Kelsh and Eisen, 2000b).

## *Genetic disorders characterised by congenital hypopigmentation in mouse and humans*

As discussed earlier the neural crest represents an important group of cells responsible for the development of many embryonic structures such as craniofacial connective tissues, sensory and autonomic ganglia, the enteric nervous system and pigmentation. Genetic disorders affecting one or more neural crest derivatives are termed neurocristopathies. By far the most commonly observed neurocristopathies are those that affect pigment development. These show a wide variety of phenotypes ranging from absence of pigmentation due to deficient melanin synthesis to abnormal distribution of pigment cells during development. On this basis it is possible to group the majority of congenital hypopigmentary disorders into 2 types.

The first type of disorders are caused by mutations in genes that code for proteins involved in the formation of the melanosome and melanin pigment at the subcellular level such as tyrosinase. Hypopigmented areas in such mutants contain melanocytes, but lack the melanin pigment.

The second type of hypomelanoses affect proteins which act at the tissue level and are responsible for migration, proliferation, survival or specification of the melanocyte lineage. Mutations in these genes result in absence of melanocytes in hypopigmented areas.

## Human disorders and their mouse models affecting melanin synthesis

Oculocutaneous albinism (OCA) represents an example of the first type of pigmentary disorders. Ocular and cutaneous melanocytes are distributed normally but are unable to synthesise melanin. OCA mutations have been classified by (King et al., 1995) depending on the location of the genetic lesion. Mutations in the gene encoding tyrosinase results in OCA1 (Barton et al., 1988). Tyrosinase catalyses the first two steps of melanin synthesis, the hydroxylation of tyrosine to dihydroxylphenylalanine (DOPA) and the oxidation of DOPA to DOPAquinone. The tyrosinase gene is the *albino* (*c*) locus of the mouse (reviewed in Oetting, 1998). In OCA2, mutations affect the P protein (murine *pink-eyed dilution* locus *p*) of unknown function, but known to be associated with the membrane matrix of melanosomes, whereas OCA3 results from genetic lesions

in the tyrosinase related protein 1, *trp-1* (Gardner et al., 1992; Boissy et al., 1996). TRP1, a DHICA oxidase, also has its functions in the melanin synthesis pathway, distal to tyrosinase. The *brown* (*b*) locus represents the murine TRP1 gene (reviewed in Oetting, 1998). Other forms of OCA such as Hermansky-Pudlak syndrome and Chediak-Higashi syndrome exhibit defects in addition to those in melanosomes (Boissy and Nordlund, 1997).

## Human disorders and their animal models affecting the distribution of melanocytes

The majority of hypomelanoses present with aberrant distribution of melanocytes and their underlying molecular defects affect various stages during melanocyte development. Animal models help to enlighten the genetic and biological basis of each of these human diseases. Table 1.1 summarises the most common human congenital pigmentary disorders and their mouse models.

### *Piebaldism and its murine homologue dominant white spotting (W)*

Piebaldism was already recognised by the ancient Romans due to its striking phenotype of often symmetrical white patches of skin and hair located on the head, chest and limbs of affected individuals (reviewed in Spritz and Ortonne, 1998). Pigmentation of the retina and irises is normal and there are no visual or hearing defects in human patients. Thus, piebaldism is thought to be a lineage-specific disorder of neural crest derived melanoblasts in skin and hair (Murphy et al., 1992; Steel et al., 1992).

Piebaldism is a rare autosomal dominant disorder (1:100 000) and inherited with equal frequencies in males and females. It is caused by mutations in *KIT*, a transmembrane type III receptor of the tyrosine kinase family (Giebel and Spritz, 1991). Signalling by this receptor requires binding of its ligand, an embryonic growth factor called Steel factor, SLF, to the extracellular ligand binding domain. Upon binding SLF, the receptor dimerises activating the intracellular tyrosine kinase domain, which then autophosphorylates specific tyrosine residues within the kinase domain. This enhances the binding of other proteins such as kinases, which mediate activation of downstream targets of the KIT-dependent signal transduction pathway (Morrison-Graham and Takahashi, 1993).

The dominant phenotypes of human piebaldism show differences in severity depending

## Summary of the most common human congenital pigmentary disorders and their mouse models

| | WS1 | WS2 | WS3 | WS4 rec | WS4 dom | HSCR2 | HSCR | Piebald | CDHS |
|---|---|---|---|---|---|---|---|---|---|
| **Gene** | *PAX3* | *MITF* | *PAX3* | *EDN3, EDNRB* | *SOX10* | *EDNRB* | *RET, GDNF* | *KIT* | *PAX3* |
| **Locus** | 2q35 | 3q12 | 2q35 | 20q13.2 | 22q13 | 13q22 | 10q11.1 | 4q12 | 2q35 |
| **mode of inheritance** | aut. dom. | aut. dom. | aut. dom. | aut. recess. | aut. dom. | aut. recess. | aut. dom. | aut. dom. | aut. dom. |
| **mouse model** | *splotch* | *mi* | *splotch* | *lethal spotting, ls* | *Dom/Dom* | *piebald lethal, $s^l$* | n/k | *white spotting, W* | *splotch* |
| **Frequency** | 1:40 000 | 1:40 000 | 1:40 000 | very rare | 1:40 000 | rare | 1:5000 | 1:100 000 | n/k |
| **M:F ratio** | equal | equal | equal | equal | 4:1 | 2:1 | 4:1 | equal | 4:1 |
| **Hypopigmentation** | √ | √ | √ | √ | √ | occ. √ | X | √ | √ |
| **heterochromic irides** | √ | √ | √ | | | occ. √ | | X | √ |
| **dystopia canthorum** | √ | X | √ | X | | X | | | √ |
| **sensorineural deafness** | √ | √ | √ | X | √ | occ. √ | | X | √ |
| **Megacolon** | rare √ | rare √ | | √ | √ | √ | extensive √ | rare √ | |
| **craniofacial defects** | mild √ | | | | | | | | more severe √ |
| **limb defects** | | | √ | | | | | | √ |

**Table 1.1:** A summary of known Waardenburg Syndrome and Hirschsprung Disease like disorders and their mouse models. Description of human phenotypes have been taken mainly from (Spritz, 1998 in Nordlund et al., 1998; Spritz and Ortonne, 1998 in Nordlund et al., 1998; Boissy and Nordlund, 1997; Pingault et al., 1997; Kuhlbrodt et al., 1998b; Pingault et al., 1998a; Southard-Smith et al., 1999). Phenotypes important in defining a particular syndrome are ticked, distinctive missing phenotypes are crossed out. Abbreviations: WS1-4, Waardenburg Syndrome 1-4; HSCR, Hirschsprung Disease; CDHS, Craniofacial-Deafness-Hand Syndrome; EDN3, Endothelin 3; EDNRB, Endothelin receptor B; GDNF, Glia derived neurothrophic factor; aut. dom./recess., autosomal dominant/recessive; mi, microthalmia; ls, lethal spotting; $s^l$, piebald lethal; Dom, Dominant megacolon, W, dominant white spotting, occ., occasional; n/k, not known; M, male; F, female.

on the site of lesion. The most severe phenotypes are observed from some missense substitutions within the conserved tyrosine kinase domain. The resulting mutant protein acts dominant negatively and thus upon dimerisation, heterozygotes have only 25% of the normal amount of functional KIT receptor dimers. In contrast, loss of function mutations caused by premature truncations of the protein in the N-terminal ligand binding domain result in only a 50% reduction of functional receptors and hence milder phenotypes due to haploinsufficiency (reviewed in Spritz, 1998).

Several animal models for piebaldism are known. Heterozygous *dominant white spotting* (*W*) and *Steel* (*Sl*) mutant mice display piebald-like phenotypes (Spritz, 1998). The *W* locus encodes the mouse homologue of the c-Kit tyrosine kinase receptor, whereas the *Steel* locus was identified as the c-Kit ligand, Steel factor (SLF). Furthermore, the zebrafish locus *sparse* (*spa*) encodes a *c-kit* homologue (Parichy et al., 1999).

All 3 animal models show depigmented areas reminiscent of the human phenotype. In addition, heterozygous *W*-mutant mice may also show hypoplastic anemia, mast cell deficiency and sterility (Spritz, 1998; Spritz and Ortonne, 1998) not yet observed in any patients or in the zebrafish mutant *spa*. Thus, mutations in *c-Kit* in mouse affect the development of melanocytes, hematopoietic precursors and primordial germ cells. Occasionally, *W* mutant mice can show abnormalities in the enteric plexus, known as Hirschsprung disease, which has also been described once in a human patient (Mahakrishnan and Srinivasan, 1980). Only one patient with a homozygous mutation in *KIT* has been reported. The child had no hair and skin pigmentation, blue irides, deafness, hypotonia, severe brachycephaly and was generally developmentally retarded (Hulten et al., 1987). This phenotype was reminiscent of viable homozygous $W^v/W^v$ mutant mice (Spritz, 1998).

By studying the zebrafish mutant allele $spa^{b5}$, a requirement of *c-kit* for melanophore migration and survival was demonstrated and confirmed previous studies in the mouse mutants (Parichy et al., 1999). $spa^{b5}$ is likely a functional null allele, since it lacks the tyrosine kinase domain. In $spa^{b5}$ mutants, a greater proportion of melanophores were observed closer to their sites of origin compared to wild-types indicating a role for *c-kit* in the migration of this cell-type. By TUNEL staining it was demonstrated that from day 4 melanophores within the epidermis apoptosed in $spa^{b5}$ mutants providing evidence for a role in survival and maintenance of the melanophore lineage. In contrast to mouse, it was shown that *spa* was not essential for melanoblast differentiation since numerous

melanoblasts were formed from the neural crest in homozygous *spa*$^{b5}$ mutants. Also, in contrast to mouse but not humans, no overt defects in primordial germ cells or hematopoietic precursors were observed in zebrafish (Parichy et al., 1999). These differences between mammalian and zebrafish c-Kit function might point to a partial functional redundancy with an additional Kit-like receptor (Parichy et al., 1999), which could have been maintained after the additional genome duplication event in teleosts (Postlethwait et al., 1998; Force et al., 1999).

### *The family of Waardenburg Syndromes*

The classical Waardenburg syndrome (WS) was first described by Waardenburg, a Dutch ophthalmologist in 1951 (Waardenburg, 1951). It is defined as an autosomal dominant disorder with piebald-like pigmentation abnormalities of skin and hair in combination with pigment abnormalities of the iris (heterochromia irides), lateral displacement of the inner canthi of the eyes, which causes a broadening of the base of the nose (dystopia canthorum) and absence or reduction of melanocytes in the cochlea (sensorineural deafness) (Spritz and Ortonne, 1998). Occasionally, patients also exhibit Hirschsprung disease, in which intrinsic ganglion cells of the myenteric (Auerbach) and submucosal (Meissner) plexi of the gastrointestinal tract are absent along a variable length of the gut. This aganglionosis causes blockage of affected regions and results in the formation of an enlarged region proximally, known as megacolon. This phenotype was named after Harald Hirschsprung, a Danish physician, who first noted a megacolon in newborn babies (Hirschsprung, 1888).
WS has since been subdivided into 4 subtypes, WS1-WS4, on the basis of the combination of these phenotypes and will be described individually.

### *Waardenburg Syndrome type I (WS1), Waardenburg Syndrome type III (WS3), Craniofacial-Deafness-Hand Syndrome (CDHS) and their mouse model Splotch (Sp)*

WS1 represents the classical WS as described above. It is a rare (1:20 000 - 1:40 000) disorder and is inherited in an autosomal dominant manner with equal frequencies in males and females (Spritz and Ortonne, 1998).
The phenotype of WS3, also called Klein-Waardenburg Syndrome, is identical to WS1, but is associated with additional musculoskeletal abnormalities of the face, limbs and upper torso (Spritz and Ortonne, 1998).

CDHS again bears strong similarities with WS1, but exhibits much more severe craniofacial abnormalities and anomalies of the hands (Asher et al., 1996).
Like WS1, both WS3 and CDHS are inherited in an autosomal dominant fashion with the same equal male and female frequencies as observed in WS1.
To date, all cases of WS1 are associated with lesions in *PAX3*, a gene encoding a paired-box domain transcription factor which additionally contains a homeobox DNA binding domain (Tassabehji et al., 1992). Surprisingly, it was shown that WS1, WS3 and CDHS are allelic and thus all 3 disorders are associated with mutations in *PAX3* (Hoth et al., 1993; Asher et al., 1996). Lesions have been identified in all 4 structural domains, the two DNA binding domains paired-box and homeobox domain, in the octapeptide and in the serine-threonine-proline-rich segment. Presumably, the loss of distinct regions of PAX3 prevents certain interactions with other transcription factors important in the protein's many roles in development. A correlation between different lesions and distinguishable phenotypes such as WS1, WS3 or CDHS could not yet be established. However, the majority of mutations in the *PAX3* gene result in WS1.

The *Splotch* (*Sp*) mouse mutants caused by lesions in *Pax3* are reminiscent of WS1 patients (Spritz, 1998). Heterozygotes of these semidominant mutations in *Pax3* exhibit white spotting in the most distal region such as the tip of the tail, and/or abdomen and feet. Homozygotes have severe neural tube defects (spina bifida and overgrowth of neural tissue) and developmental abnormalities in several neural crest derived structures such as total lack of body pigmentation and abnormalities in spinal ganglia, DRG, sympathetic and enteric ganglia (Auerbach, 1954; Lang et al., 2000). Sp/Sp mutant mice die in utero on E13 (Russell, 1947). Only one child with apparently homozygous WS1 has been reported. Its symptoms were more severe than the heterozygous phenotype including dysmorphic facial features, almost completely white hair and extensive loss of skin pigmentation (Zlotogora et al., 1995).
*Pax3* was shown to be expressed early in the embryonic primitive streak, differentiating dorsal neuroepithelium, neural crest derived structures and regions of the adult brain (Goulding et al., 1991; Goulding et al., 1993). Recently, *Pax3* was proposed to be required early during neural crest development to expand a pool of fate-restricted progenitor neural crest cells such as melanoblasts, but not for their migration (Hornyak et al., 2001). Thus, even in homozygous *Sp/Sp* mutants, melanocytes and other neural crest derived neural tissues were observed in characteristic locations, but in fewer

numbers. The phenotype of WS1, WS3 and CDHS patients is also consistent with an early role for PAX3 affecting the proliferation of several neural crest derivatives.

### *Waardenburg Syndrome type II (WS2)*

Waardenburg Syndrome type II is an autosomal dominant disorder that affects approximately 1:20 000 – 1:40 000 newborns with equal frequencies in males and females. Clinically, the disease is vaguely defined with patients presenting a similar phenotype to WS1, but lack dystopia canthorum (reviewed in Spritz, 1998). Pigmentation defects in skin, hair and eyes tend to be less severe than in WS1 and hearing loss due to a lack of melanocytes in the cochlea is not always present (Tassabehji et al., 1994).

At least in some individuals with WS2, a mutation in *MITF*, a basic helix-loop-helix-leucine zipper protein, has been identified. Studies on mutant mouse (*microphthalmia*) and zebrafish (*nacre*) *MITF* homologues have been crucial in elucidating possible biological mechanisms underlying WS2.
Heterozygous *microphthalmia* (*mi*) mice show white spotting on the belly, head and tail and reduced pigmentation in the iris. Homozygotes on the other hand have small unpigmented eyes, lack all melanocytes in the skin and in the inner ear, rendering them deaf. Some strong *mi* alleles additionally have a deficiency in retinal pigment cells, in mast cells and deficiency in secondary bone resorption (osteopetrosis) (Steingrimsson et al., 1994). The original *mi* allele derived from an irradiated male shows semidominance and a strong phenotype (Hodgkinson et al., 1993). Other alleles such as *eyeless-white* ($mi^{ew}$) and *cloudy-eyed* ($mi^{ce}$) are recessive and *spotted* ($mi^{sp}$) has a normal appearance (Steingrimsson et al., 1994).
Mi, typically for basic helix-loop-helix-leucine zipper proteins, binds DNA through the basic domain and forms homo-and heterodimers utilising the helix-loop-helix and the leucine zipper domain (reviewed in Kadesch, 1993). This mode of action also helps to explain the differences reported for the various alleles, whereby a mutation severely affecting the protein's function, but retaining its ability to dimerise causes a dominant or semidominant appearance due to dominant negative action. A recessive allele is predicted to encode a loss of function allele, e.g. a prematurely truncated protein or one with a non-functional dimerisation domain (Steingrimsson et al., 1994).
Surprisingly, the human alleles in family WS.026 and WS.002 show dominance although the mutations affect splice sites in the first few exon boundaries and very

severely truncate the protein. Surely, this protein then should act as a null allele and thus be recessive as observed in various mouse *mi* alleles (Tassabehji et al., 1994). Two possible explanations have been proposed. The *MITF* gene dosage might be more critical in humans than in mouse (haploinsufficiency) or perhaps the altered splice site could lead to exon skipping and thus to a protein that can still dimerise and act in a dominant negative manner (Tassabehji et al., 1994).

During mouse development, *mitf* is expressed in presumptive melanoblasts contributing later to skin pigmentation and melanocytes in the stria vascularis of the inner ear. Furthermore, expression was observed in the embryonic heart and there was weak staining in the outer retinal layer of the eye. In adults, *mitf* was highly expressed in the heart and at lower levels in skeletal muscle (Hodgkinson et al., 1993). Mitf acts cell autonomously, consistent with coding for a transcription factor. It was shown to be required for committed melanoblast survival in premigratory neural crest cells perhaps by inducing a set of trophic genes (Hornyak et al., 2001). Later during melanocyte differentiation Mitf is thought to activate and regulate pigmentation gene expression by binding to M-box elements in the promoter of target genes. Such M-box elements were found in at least 3 enzymes involved in pigmentation, tyrosinase, tyrosinase-related protein 1 and 2, and in *in vitro* studies, Mitf was able to transcriptionally activate a reporter gene driven by an M-box element in conjunction with a TATA box of a pigmentation gene promoter (Hemesath et al., 1994).

Fairly recently, the zebrafish orthologue of *mitf* has been cloned, *nacre* (*nac*), and was shown to be highly important in melanophore development (Lister et al., 1999). Homozygous $nac^{w2}$ fish lack neural crest derived melanophores throughout embryonic development. In adults, the pigmentation of the retinal epithelium is normal, but the number of xanthophores is slightly reduced and the number of iridophores is increased by approximately 40% at day 3. It was shown that *nacre* functioned early within the pigment lineage since early melanoblast markers such as *tyrosinase-related protein 2* (*trp2*), now called *dopachrome tautomerase* (*dct*; Kelsh et al., 2000c), and *spa* were almost or entirely absent from neural crest cells. Wild-type neural crest cells transplanted into $nac^{w2}$ hosts differentiated into melanophores with wild-type morphology indicating cell-autonomy for *nacre* within this lineage consistent with mouse *mi*. No overt defects in any other neural crest derivatives such as neurons and glia of the peripheral nervous system and craniofacial structures were detected. Interestingly, it was demonstrated that *nacre* was able to induce ectopic pigment cells of

abnormal morphology when misexpressed. This pigmentation was observed 5 hours earlier than in wild-types and in cells of non-crest origin since expression of *dct* in those cells was apparent before the onset of neural crest cell migration. These findings support a role for *nacre* as a melanogenic factor activating the differentiation program possibly in cooperation with additional transcription factors (Lister et al., 1999).
The lack of the retinal phenotype in zebrafish observed in humans and mice could be explained by the identification of a second *mitf* homologue. Mitfb was shown to act redundantly with Mitfa (Nacre) in the retinal pigment epithelium (RPE; Lister et al., 2001). Mitfb can also explain the recessive phenotype in *nacre* embryos, since it was able to induce ectopic melanophores in *nacre* embryos.

In summary, both mouse and fish *mitf* mutants model the human phenotype of WS2 very well. During development, Mitf appears to have an early role in melanoblast survival and later in differentiating melanocytes by regulating the transcription of pigmentation genes such as tyrosinase. *MITF* itself was shown to be directly regulated by PAX3, which would explain the similar phenotypes of WS2 and WS1/WS3. It thus suggests that the defects observed in at least some WS1 patients might partially result from the inability of PAX3 to regulate *MITF* (Watanabe et al., 1998).

### *Recessive Waardenburg Syndrome type 4 (WS4)*

Waardenburg syndrome type 4, also known as Waardenburg-Shah Syndrome, combines certain phenotypes exhibited in WS3 with Hirschsprung disease, although there is no deafness and sometimes the extent of hypopigmentation is more extreme (Gross et al., 1995). Shah and his colleagues reported studies on 12 babies all of whom had white forelocks, eyebrows and eyelashes and intestinal obstruction. At least 8 exhibited normal pigmentation of the eyes and at least 6 of them lacked dystopia canthorum or white skin patches. None of them had detectable deafness, all of them died 3-38 days after birth and the locus seemed inherited in an autosomal recessive manner (Shah et al., 1981). Lesions in the human *endothelin-3* (*EDN3*) gene have now been discovered in patients suffering from the recessively inherited WS4 (Edery et al., 1996; Hofstra et al., 1996). EDN3 is a 21 amino acid peptide that binds to the endothelin receptor B (EDNRB).

The natural mouse mutation, *lethal spotting* (*ls*), was mapped to the *Edn3* locus and serves as a useful animal model for WS4. The *ls* phenotype is caused by a point

mutation that prevents the conversion of preproendothelin-3 to its functional form (Baynash et al., 1994).

Furthermore, targeted disruption of the *endothelin-3* (*Edn3*) gene in mice also gave rise to recessively inherited spotted coat colour and Hirschsprung disease reminiscent of the human WS4 phenotype. Homozygous *edn3/edn3* mutant mice have black eyes and dark coat patches on their head and hips, but 70-80% of their coats lack melanocytes. Their ileum and cecum is largely distended, the entire colon is aganglionic and only approximately 15% of homozygous mice survive to adulthood.

Studies indicated that EDN3 promoted survival and differentiation of melanoblasts in neural tube cultures (Lahav et al., 1996).

Transgenic experiments were employed to enlighten the aganglionic phenotype in mice. To label enteric neuroblasts, the dopamine-beta-hydroxylase (DBH) promoter was used to drive lacZ expression in this cell-type, but enteric neuroblasts were never observed in the distal colon of *ls/ls* embryos (Kapur et al., 1992; Kapur et al., 1993). Retarded colonisation was first evident at the junction of the small and large intestine. This suggests a pancolonic defect in *ls/ls* mice and thus an important role for Edn3 to ensure complete colonisation of the gut (Kapur et al., 1993).

### *Dominant Waardenburg Syndrome type 4/Hirschsprung Disease*

This autosomal dominant form of WS4 affects approximately 1:40 000 newborns with approximately 4 times higher frequencies in males than in females. WS4 patients present with pigmentation anomalies of skin and hair and sensorineural deafness, similar to WS1. In addition WS4 patients suffer from aganglionic megacolon, but usually lack dystopia canthorum and musculoskeletal malformations (Pingault et al., 1998a; Southard-Smith et al., 1999).

The mouse mutant *dominant megacolon, Dom*, arose spontaneously in the Mouse Mutant Stock Centre of the Jackson Laboratory, Bar Harbor, ME (Lane and Liu, 1984). It is a particularly interesting model for WS4 as it mimics well the variable penetrance and expressivity of the megacolon phenotype in affected patients (Southard-Smith et al., 1999). The heterozygous *Dom/+* phenotype, similar to homozygous lesions in *Edn3* (*ls/ls mice)* and *Ednrb* ($s^l/s^l$ mice), has been described as white belly spot, white feet and deficiencies of myenteric ganglion cells. By embryonic day 5 (E5) pigmentation anomalies become apparent and by E10 heterozygous mice appear malnourished whereas their abdomens look distended due to underlying megacolon

(Lane and Liu, 1984). Hearing loss in mice is difficult to assess, but circling and hyperactivity was observed (Pingault et al., 1998a). This behaviour is known as waltzing syndrome and suggestive of inner ear defects (Deol et al., 1986). The life span of *Dom*/+ mice usually ranges from 2 days to 12 months depending on the genetic background. Homozygous *Dom/Dom* mice generally exhibit a more severe phenotype and die in utero around E13 (Lane and Liu, 1984).

*Dom* was mapped to chromosome 15 (Lane and Liu, 1984) between microsatellite markers *D15Mit68* and *D15Mit2* (Pingault et al., 1997) and found to encode Sox10, an HMG domain type transcription factor (Herbarth et al., 1998; Southard-Smith et al., 1998). In mice, *Sox10* expression is first detected at E8.5 (Southard-Smith et al., 1999), is found in newly forming cranial and trunk neural crest cells in wild-type and $Sox10^{Dom/Dom}$ mice by E9.5 and in melanoblasts on the lateral pathway and all cranial, dorsal root, sympathetic and enteric ganglia later in embryonic development. From E9.5 to E12.5, *Sox10* expression is observed in the epithelium of the otic vesicle, later in the developing cochlea and vestibule and in the organ of Corti in adult mice. *Sox10* expression levels generally decrease over time and by E13.5 migrating melanoblasts are *Sox10*-negative including those migrating into the stria vascularis of the inner ear (Watanabe et al., 2000).

In homozygous $Sox10^{Dom/Dom}$ mice *Sox10* levels are drastically reduced in cranial ganglia and nerves at E10.5 and all *Sox10* expression is lost by E11.5 apart from in the most caudal neural crest. Thus, in contrast to wild-types, no *Sox10*-positive cells can be detected in homozygote gut sections at E14.5 (Herbarth et al., 1998; Southard-Smith et al., 1998). Heterozygous $Sox10^{Dom/+}$ mice exhibit the same expression patterns as wild-types, but levels are overall reduced. Enteric neuron number is decreased and colonisation is somewhat retarded from E11.0 over the entire length of the gut, not just in the large intestine as described for *ls/ls* and $s^l/s^l$ mice. Even so, growth and rotation of the intestine proceeds normally (Kapur et al., 1996). In one study, 14 out of 22 $Sox10^{Dom/+}$ mice had extensive aganglionosis at E14 varying between being localised only in the distal part of the rectum and as proximal as the ileocaecal junction (Puliti et al., 1996). In Northern blot hybridisations *Sox10* transcripts were detected in heart, brain, lung, skeletal muscle and testes in adult mice (Southard-Smith et al., 1999).

After *Sox10* was shown to cause the *Dom* phenotype, heterozygous lesions were also detected in the human *SOX10* homologue in previously unidentified WS4 patients (Pingault et al., 1998a; Southard-Smith et al., 1999). Like in mice, patients also show a

variable phenotype even within a family, which has been hypothesised to be due to modifier loci. Mice show a fully penetrant phenotype on a mixed C3HHeB:FeJLe-axC57BL/6JLeB6 (C3xB6) background, whereas lethality is higher (70% versus 18%) and head spotting is lower (3% versus 50%) on a B6 background compared to a C3 background (Southard-Smith et al., 1999). Similarly, all *Sox10*$^{Dom/+}$ mice completely lack enteric ganglion cells in the distal large intestine on a B6 background in comparison to only one third on C3 (Kapur et al., 1996). One such locus modifying the pigmentation and loss of hearing phenotype in humans has been mapped to chromosome 21 (Puffenberger et al., 1994a). A similar modifier locus has been found on mouse chromosome 10, which segregates with the mast cell growth factor (MGF), a c-Kit ligand, and thus makes a good candidate gene (Southard-Smith et al., 1999).

In human embryos, *SOX10* is expressed in the neural crest at 4 weeks, in all cranial ganglia, DRG, cranial and spinal nerves, the otic vesicle and in the entire sympathetic and parasympathetic ganglion chain by 5-6 weeks. At 6 weeks of age, also enteric ganglia, cranial foregut, oesophagus, stomach and lungs, but not yet the CNS are *SOX10*-positive. Later on, at 17-25 weeks, *SOX10* is mostly observed in several parts of the brain, in lung and kidneys (Bondurand et al., 1998a). In adults, transcripts were detected in Northern Blot hybridisations in the heart, brain, colon, small intestine, lung, skeletal muscle, testes, ovary, pancreas, bladder, prostate, stomach, spinal cord, trachea and adrenal gland (Bondurand et al., 1998a; Southard-Smith et al., 1999). These high transcript levels in brain and spinal cord in humans is consistent with *Sox10* expression observed in rodent glial cells (Kuhlbrodt et al., 1998a). Also, like in mice, expression seen in the entire intestinal tract, stomach, small intestine and colon corresponds to the developing enteric nervous system (Southard-Smith et al., 1999).

Mouse and human expression patterns are very similar albeit not identical. Only humans show expression in the cephalic neural crest giving rise to cartilaginous structures in nasal bones, in the fetal and adult cerebral cortex, in major brain nuclei, heart, prostate and testis (Bondurand et al., 1998a).

Inoue and colleagues recently identified a novel *SOX10* mutant allele in a patient with myelination deficiencies in the CNS and PNS in addition to hypopigmentation, long-segment HSCR, heterochromic irides, deafness, dystopia canthorum and nystagmus (Inoue et al., 1999). This finding was quite surprising since long-segment HSCR and facial malformations are usually associated with mutations in *RET* and *PAX3*, respectively. Also, such myelination defects have not yet been reported for *Sox10* in any

model organism (Inoue et al., 1999). They can be accredited to a role for Sox10 in myelin development and maintenance though since *Sox10* is expressed in Schwann cells and oligodendrocytes known to produce myelin. Sox10 also interacts with myelin-related transcription factors such as Oct6, Pax3 and EGR2 (Kuhlbrodt et al., 1998a). Furthermore, a child diagnosed with Yemenite Deaf-Blind Syndrome was found to exhibit a heterozygous *SOX10* mutation. The girl presented patches of hypo-and hyperpigmentation on trunk and extremities, grey hair, white eyebrows and eyelashes, severe hearing loss, nystagmus and dental abnormalities, but no dystopia canthorum or HSCR (Bondurand et al., 1999).

Several roles for Sox10 during neural crest development have been suggested such as involvement in survival, migration and differentiation. Recent studies have demonstrated synergistic activation of several target genes by Sox10 with partner factors. For example, Pax3 and Sox10 bind to enhancer elements in the *cRet* (Lang et al., 2000) and in the *Mitf* promoter (Bondurand et al., 2000; Potterf et al., 2000) and Sox10 also interacts synergistically with Tst-1/Oct6/SCIP (Kuhlbrodt et al., 1998a). However, further studies will be required to fully elucidate Sox10's complex involvement during neural crest development.

### *Hirschsprung Disease/Waardenburg Syndrome (HSCR2)*

Puffenberger and his colleagues studied a large Mennonite pedigree in which 6.3% of affected individuals had bicoloured irides, 2.5% were hypopigmented, 5.1% showed hearing loss and 7.6% had a white forelock additional to the HSCR phenotype. This autosomal recessive disease was named Hirschsprung Disease/Waardenburg Syndrome or Hirschsprung disease type 2 (HSCR2) and mapped to chromosome 13q22, as did *endothelin receptor B* (*EDNRB)* encoding a G protein-coupled receptor (Puffenberger et al., 1994a; Puffenberger et al., 1994b).

HSCR2 is distinguished from WS4 by the presence of bicoloured irides. However, in contrast to WS1 and WS3, dystopia canthorum is absent in both WS4 and HSCR2. It was estimated that approximately 95% of HSCR2 caused by mutations in EDNRB resulted in short-segment aganglionosis (Chakravarti, 1996).

Its mouse model, the *piebald-lethal* ($s^l$) mouse, also exhibits a white colour coat in addition to megacolon (Hosoda et al., 1994). The hypomorphic allele *piebald* (*s*) only shows 20% white coat spotting and almost never enteric aganglionosis even in homozygous *s/s* mutants. Heterozygous $s^l$ mice appear to be normal, whereas in the

homozygous state, over 90% of the coat pigmentation is lost apart from very small dark asymmetrical regions on the head and hips, eyes are black and pups never survive to adulthood. Abnormalities in the inner ear were also described (Hosoda et al., 1994). The intestinal region distal to the sigmoid colon usually lacks myenteric ganglion neurons causing distension and sometimes even perforations in more proximal regions. The melanoblast marker *trp-2* (*dct*) is lost by E10.5 in *Ednrb/Ednrb* mice, whereas it is still present in *W* and *Sl* mutant mice at E11 indicating that EDNRB may function upstream of *c-kit* and *Steel factor* in melanoblast development (Hosoda et al., 1994). Gariepy and colleagues (1998) were able to show that EDNRB was needed in the enterics lineage after the latter had segregated from the melanocyte precursors. With transgenic $s^l/s^l$ rats containing the *Ednrb* gene under the control of a human dopamine-beta-hydroxylase (DBH) promoter, which directs *Ednrb* expression to the enteric nervous system precursors, the megacolon phenotype was rescued, but not the coat pigmentation defect. Kapur demonstrated in $s^l/s^l \leftrightarrow$ wild-type aggregation chimeras that EDNRB although being a receptor had a non-cell autonomous effect on melanoblasts and neuroblasts indicating that inter-cellular signals downstream of EDNRB influence colonisation. Hence, wild-type neuroblasts were able to rescue migration of $s^l$ mutant neural crest cells in some chimeric embryos (Kapur et al., 1995).
Shin and his colleagues could further show a temporal requirement of EDNRB between E10 and E12.5 for the migration of both melanoblasts and enteric neuroblasts (Shin et al., 1999).

Although the severity of defects in *Ednrb* and *Edn3* mutant mice are similar as one might expect since they act as receptor and its ligand during development it is interesting to note that the *Edn3* phenotype appears to be milder. This could be explained by slight redundancy with closely related endothelins, Edn1 and Edn2, which are known to be able to bind to EDNRB albeit with lower affinity (Hosoda et al., 1994).

Lethal white foal syndrome (LWFS) also represents a homologous congenital disorder to HSCR2 in horses. In its homozygous state it is characterised by white coat colour and aganglionosis of the bowel. Heterozygotes exhibit a specific coat colour pattern known as *overo*, but no intestinal defects. The location of the mutant lesion was placed in the *EDNRB* gene (Yang et al., 1998).

### *Dominant Hirschsprung Disease without pigmentation defects*

An autosomal dominantly inherited form of Hirschsprung disease was found to be caused by mutant lesions in the proto-oncogene *RET* (**RE**arranged during **T**ransfection; Edery, Lyonnet et al. 1994; Romeo, Ronchetto et al. 1994). Apart from Hirschsprung disease, RET, a receptor tyrosine kinase, has previously been implicated in multiple endocrine neoplasia type 2A and B (MEN2A, B) and medullary thyroid carcinoma, but never in pigmentation anomalies (Spritz, 1998). It was estimated that mutations in *RET* account for approximately 50% of HSCR cases, of which 75% involve long-segment aganglionosis (Chakravarti, 1996). The latter is defined by the extent of the aganglionotic tract reaching proximal to the splenic flexure (Romeo et al., 1994).

In zebrafish, the orthologue of *c-ret* is expressed in spinal motoneurons, pronephric ducts, cranial ganglia, pharyngeal arches and the enteric nervous system (Bisgrove et al., 1997).

During mouse development, *c-ret* is expressed in sensory, autonomic and enteric ganglia, the Wolffian duct and in the ureteric bud epithelium. Mice with a targeted homozygous mutation in *c-ret* exhibit defects in kidney development, total aganglionosis posterior to the proximal stomach, lack the superior cervical ganglia (SCG) and die soon after birth (Romeo et al., 1994; Schuchardt et al., 1994). In contrast to humans, *c-ret* in mice behaves in an autosomal recessive manner.
Homozygous targeted mutations in one of c-Ret's ligands, Glial cell-line derived neurotrophic factor (GDNF), results in a very similar and equally severe phenotype (Schuchardt et al., 1994). It was demonstrated *in vitro* that RET signalling via GDNF is necessary and sufficient (together with the GPI-linked co-receptors) for development of cultured neuronal cells, whereby GDNF promotes survival, proliferation and differentiation of *c-ret* positive enteric neural crest cells during E12.5-E13.5. This is consistent with increased programmed cell death observed by TUNEL staining in enteric neural crest cells in homozygous *c-ret* mutant mice (Taraviras et al., 1999). Mutations in *Pax3*, *Sox10* and *c-ret* can result in an aganglionic phenotype. Consistent with this, results indicated that Pax3 and Sox10 together might be required to activate a *c-ret* enhancer element (Lang et al., 2000). Thus, the failure of dysfunctional Sox10 or Pax3 to activate *c-ret* expression sufficiently might account for the enteric phenotype observed in these mutants.

## *Sox10 as a candidate gene for the colourless mutation*

As described in the previous section, there are numerous neurocristopathies reported in mice and humans affecting pigmentation in combination with defects in other neural crest derivatives (Table 1.1). Most of those loci present with similar combinations of phenotypes to *cls* embryos, but can be excluded as candidate genes for *cls* after more careful analysis of their defects.

The extent and severity of aganglionosis in *cls* embryos is most reminiscent of *c-ret* mutations in mice. Like *cls*, *c-ret* behaves in an autosomal recessive manner and homozygous mice exhibit total aganglionosis posterior to the proximal stomach (Romeo et al., 1994; Schuchardt et al., 1994). In humans, *RET* mutations are dominant, but still present long-segment aganglionosis in 75% of all cases (Chakravarti, 1996). However, neither in mice nor humans pigmentation anomalies are associated with the *c-ret/RET* locus and this makes *RET* an unlikely candidate for *cls*.

The *PAX3* locus associated with WS1, WS3 and CDHS can be excluded for similar reasons. All three human disorders involve craniofacial defects like dystopia canthorum (reviewed in Spritz, 1998), whereas ectomesenchymal derivatives are unaffected in *cls* embryos. Furthermore, megacolon associated with aganglionosis is only very rarely observed in patients with *PAX3* mutations and is not present in the one patient reported with homozygous WS1 (Zlotogora et al., 1995) or homozygous *splotch* mice (Russell, 1947; Auerbach, 1954).

Likewise, mutations in *KIT* and *MITF* mostly cause pigmentation defects and again, megacolon is only very rarely observed (Mahakrishnan and Srinivasan, 1980). Although the phenotype of homozygous mice and humans is aggravated, there are no reports of aganglionosis.

Mutations in *EDN3* and *EDNRB* are both recessively inherited in mice and men and combine defects in pigmentation and enteric nervous system. The homozygous enteric ganglion phenotype appears to be milder compared to homozygous *cls* embryos, since even *s/s* and *ls/ls* mutant mice only lack enteric ganglia in the posterior gut (Kapur et al., 1992; Kapur et al., 1993; Hosoda et al., 1994). However, this could be caused by the genetic background or species specific differences.

*cls* embryos exhibit a grossly abnormal ear morphology (Kelsh et al., 1996; Malicki et al., 1996; Whitfield et al., 1996). However, deafness has not been reported in WS4

patients with lesions in *EDN3* and only in 5.1% of HSCR2 patients suffering from lesions in *EDNRB* (Puffenberger et al., 1994a).
Furthermore, *Ednrb/Ednrb* mouse mutants lack the extensive defects in the peripheral nervous system such as cranial ganglia and DRG observed in *cls* embryos (Southard-Smith et al., 1998).

The pigmentation defect in *cls* embryos is caused by a severely reduced number of melanoblasts. Similarly, in homozygous $Sox10^{Dom/Dom}$ mice melanoblasts are absent from E10.5 (Southard-Smith et al., 1998). The extent and severity of the enteric ganglia phenotype of *cls* mutants is strongly reminiscent of homozygous $Sox10^{Dom/Dom}$ mice. The latter exhibit extensive aganglionosis over the entire length of the gut by E14.5 (Southard-Smith et al., 1998). Furthermore, $Sox10^{Dom/Dom}$ mice and *cls* embryos have extensive peripheral nervous system defects outside the enteric lineage, for example in cranial ganglia and DRG (Herbarth et al., 1998; Southard-Smith et al., 1998).
*cls* embryos exhibit inner ear defects, since only a single sensory patch and 2 small otoliths are formed (Whitfield et al., 1996). In mice, *Sox10* is expressed in the developing ear and although hearing loss has not yet been conclusively demonstrated, circling and hyperactivity has been observed in $Sox10^{Dom}$ heterozygotes suggesting inner ear defects (Pingault et al., 1998b). Lastly, ectomesenchymal neural crest derivatives contributing to the craniofacial skeleton for example are unaffected in both *cls* and $Sox10^{Dom}$ mutants.

In summary, of all the known pigmentation loci, homozygous *cls* embryos appear to be most reminiscent of homozygous $Sox10^{Dom/Dom}$ mouse mutants and thus, *Sox10* was chosen as the most promising candidate gene for the zebrafish *colourless* mutation.

## *The Sox gene family of transcription factors and their role in development*

There are a number of different groups of DNA binding proteins, each of which are characterised by their DNA binding domain. The latter consists of conserved structural motifs such as the basic helix-loop-helix motif, leucine zippers or zinc fingers. In 1990, Tjian and his coworkers (Jantzen et al., 1990) discovered a novel DNA binding domain in the gene for the RNA polymerase I transcription factor UBF. This region is repeated six times in UBF and each of them shows high homology to two domains of High

Mobility Group1 (HMG1) proteins, which gave this new binding domain its name. Genes that were found to contain such an HMG domain were classified as members of the HMG box superfamily. Since Tjian and his colleagues first recognised the HMG box as a binding domain, many more members have been isolated from organisms as diverse as mammals, insects, plants and fungi. This indicates that a putative ancestral gene giving rise to the HMG box superfamily predates the separation of animal and fungal kingdoms at least 1000 million years ago (Laudet et al., 1993).

## Separation of HMG box proteins into two major subfamilies based on the properties of their DNA binding domain

Several groups have attempted to partition this superfamily into smaller subfamilies by comparing the identity of the most conserved HMG domain between members. On average, the HMG domains are comprised of 80 amino acids and show approximately 25% sequence identity between different homologues. There are only 3 residues conserved in >80% of all members, P at position 8, W at position 45 and K at position 53 of the alignment published by (Laudet et al., 1993). Based on this sequence alignment an unrooted phylogenetic Fitch tree was calculated dividing the HMG box superfamily into two distinct subfamilies, the HMG/UBF and the TCF/SOX subfamily.

Members of the HMG/UBF subfamily such as HMG-1, UBF-1, ABF2, TETHMG and CCG1 contain two or more HMG domains (Laudet et al., 1993) and references therein). HMG-1 like proteins are ubiquitously present in all eukaryotes and roles have been suggested in DNA replication and nucleosome assembly. They bind DNA in a manner that is independent of sequence, but instead recognise certain DNA structures *in vitro* like four-way junction DNA and $(CA)_n$ repeats forming the latter (Ferrari et al., 1992; Peters et al., 1995). Such four-way junction DNA structures can be generated *in vivo* by recombination events and by intrastrand base pairing of inverted repeats (Bianchi et al., 1989).

The TCF/SOX subfamily includes transcription factors like TCF-1, LEF-1, MATA1 and SRY-like *Sox* genes (Laudet et al., 1993 and references therein). Interestingly, with their single HMG domains they recognise both particular DNA structures like four-way junctions and also specific heptameric consensus sequences on target genes based on the motif 5' (A/T A/T C A A A/T G) 3' (Harley et al., 1994). The human SRY binding domain was reported to exhibit a greater affinity for junction DNA

*in vitro* than for the specific DNA binding site. The interaction between SRY and DNA which is already distorted, closely mimics the final structure that is formed after specific binding and conformational bending of the DNA, but is much less costly energetically. However, once bound to the consensus motif, the interaction between DNA and protein is much stronger. Ferrari and her coworkers proposed a model to explain how this subfamily of HMG domain proteins were still able to act through sequence specificity even though binding to kinked unspecific DNA is favoured. They suggested that the concentration of HMG1-like proteins recognising distorted DNA with great affinity was much higher than of transcription factors like SRY and so would bind with higher probability. Statistically, SRY would get displaced from these sites and restricted to its DNA specific, undistorted binding sites, which HMG1-like proteins were unable to bind (Ferrari et al., 1992).

The 3 dimensional structure of the HMG domain with and without being bound to consensus target DNA was solved with the use of NMR by 2 separate groups (van Houte et al., 1995; Werner et al., 1995). Irrespective of the presence of DNA the domain forms three α-helices, with helix I antiparallel to helix II forming the long arm of a twisted L-shape, and with helix III forming the short arm. A hydrophobic core supports this concave structure.

In the TCF/SOX subfamily this L-shape recognises the heptameric consensus motif as described above and binds to the minor groove of target DNA. This is an unusual mode of action as most other transcription factors bind to the major groove of DNA. As a result, a bend of up to 130° is induced in the target DNA which slightly unwinds the helix and thus functions to organise the local chromatin structure (Pevny and LovellBadge, 1997). Furthermore, close proximity to other transcription factors that bind to the major groove becomes physically possible and facilitates the formation of transcriptionally active multiprotein complexes (Wolffe, 1994; Werner and Burley, 1997). Because of their ability to bend DNA, members of this gene family have also been known as architectural transcription factors (Grosschedl et al., 1994). To initiate the widening of the minor groove, it was demonstrated that Ile-168 in human SRY (can be replaced by Met or Phe in other sequence-specific HMG box proteins) disrupts the base pair stacking at the site of insertion by intercalating between two AT basepairs. (Love et al., 1995; Peters et al., 1995; Werner et al., 1995). In one of the many human 46X,Y sex reversal mutations, Ile-168 is replaced by Thr resulting in a diminished DNA

binding affinity by two orders of magnitude compared to wild-type SRY (Werner et al., 1995).

## Subdivision of the SOX family based on overall sequence and structural homologies

Within the TCF/SOX family, the SRY-type SOX proteins are the best characterised group of transcription factors. Until recently, a SOX protein was distinguished from TCF/LEF proteins by its high amino acid identity (>50%) within its HMG domain to that of the founding member of this group, mouse SRY. Thus, genes containing such an SRY-type HMG box were called *Sox* genes (Pevny and LovellBadge, 1997).

Koopman and his colleagues have proposed a new method to identify members of the SOX protein subfamily (Bowles et al., 2000). They suggest using the conservation of key motifs within the HMG domain as an alternative criterion to define *Sox* genes. They argue that SRY is not the most suitable representative for the SOX protein family as it is the only member which has no homologues outside the mammalian lineage. Additionally, most likely because of their location on the Y chromosome, the rate of evolution for SRY homologues is a lot faster and hence the subgroup of SRY genes is clearly more divergent than others . If the homology to the SRY HMG domain was to be used to identify SOX proteins, >50% amino acid identity was a too stringent criterion. Their argument is backed up by the recently isolated huSOX30 and ceSOXJ proteins, which only show 48% and 46% sequence identity to hu SRY, respectively. However, their latest alignment of HMG domain sequences revealed that the motif RPMNAFMVW is common to all non-SRY SOX protein members and they suggest this as the most reliable signature to distinguish SOX proteins in future (Bowles et al., 2000).

SOX proteins were thus divided into 10 subgroups A-J (Figure 1.1). The subdivisions based on the phylogenetic alignment of HMG domain sequences are consistent with the grouping of SOX proteins based on overall structural and functional motifs of the entire protein sequence. Such structural motifs include the conservation of intron positions, whereas functional motifs refer to additional domains characteristic for the subgroup. Members of the subgroup D for example (Sox5, Sox6, Sox13 and Sox23), are characterised by a leucine zipper domain, which was shown to be capable of homodimerization (Lefebvre et al., 1998). This dimerization only indirectly interferes

with DNA binding in that it reduces the affinity of the dimer for single Sox consensus binding sites. For the long variant of Sox5,L-Sox5 heterodimerization with Sox6 has also been demonstrated (Lefebvre et al., 1998). Whereas most members of subgroups B1 (Sox1, Sox2, Sox3 and Sox19), C (Sox4, Sox11, Sox22 and Sox24), E (Sox8, Sox9 and Sox10) and F (Sox7, Sox17 and Sox18) have one or two transactivation domains, subgroup B2 proteins (Sox14 and Sox21) contain a repression domain (Uchikawa et al., 1999; Bowles et al., 2000).

Sox proteins exhibit very high sequence conservation within the DNA binding domain, especially between members of the same subgroup. Furthermore, Sox proteins often show overlapping regions of expression and thus, the question of how they can still regulate the transcription of target genes in a cell-type specific manner arises.

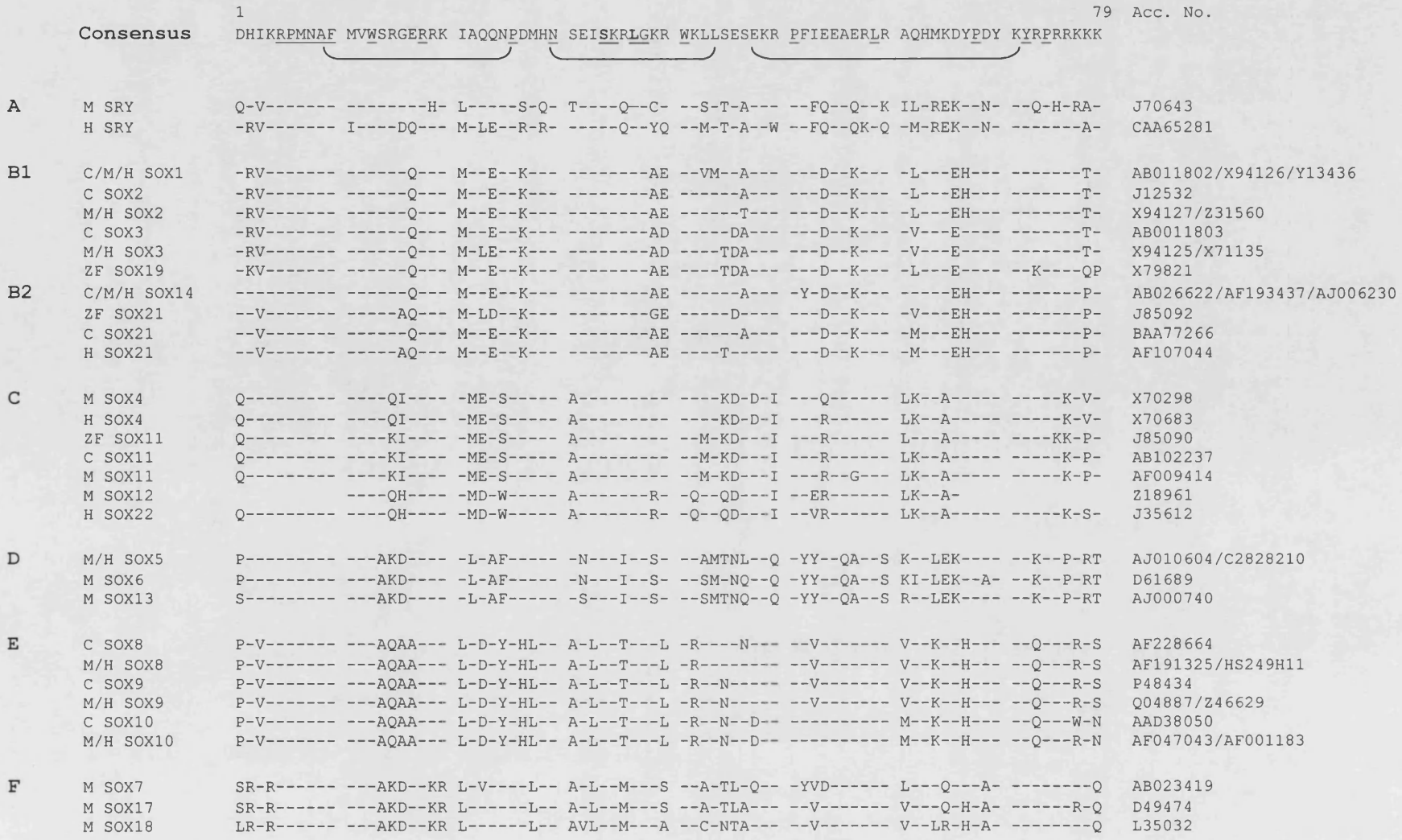
## How Sox genes exhibit cell-specific transcriptional regulation and their role in development

Several mechanisms might be expected to contribute to this specificity such as additional specific flanking nucleotides to each side of the consensus binding motif and other functional domains in addition to the DNA binding domain. Furthermore, synergistic activity with other cell-type specific transcription factors might increase the specificity dramatically in some cases.

The recognition motifs to which SOX proteins have been shown to bind vary very little in sequence both *in vivo* and *in vitro* (Kamachi et al., 1999; Mertin et al., 1999). The binding affinity might be increased slightly by specific flanking nucleotides to each side of the core element AACAAT, as was demonstrated for the human SOX9 and SRY protein *in vivo* (Mertin et al., 1999).

The angle SOX proteins introduce by binding to DNA depends largely on the binding site and thus is also fairly similar. Hence, the specificity cannot be explained solely by selective binding to downstream targets or by introducing the correct degree of bending. For example, both HMG domains of SOX1 and SOX9 from chicken were able to bind equally well to the Sox binding site in the Col2a1 minimal enhancer element, which is only a target of SOX9 *in vivo* (Kamachi et al., 1999).

There are indication that other domains might contribute significantly to specificity apart from the DNA binding domain and the C-terminal transactivation domain. Cell-type specificity of SOX1, SOX2 and SOX9 is observed *in vivo*, despite of



| | | 1 | 79 | Acc. No. |
|---|---|---|---|---|
| | Consensus | DHIKRPMNAF MVWSRGERRK IAQQNPDMHN SEISKRLGKR WKLLSESEKR PFIEEAERLR AQHMKDYPDY KYRPRRKKK | | |
| A | M SRY | Q-V------- --------H- L-----S-Q- T----Q--C- --S-T-A--- --FQ--Q--K IL-REK--N- --Q-H-RA- | | J70643 |
| | H SRY | -RV------- I----DQ--- M-LE--R-R- -----Q--YQ --M-T-A--W --FQ--QK-Q -M-REK--N- -------A- | | CAA65281 |
| B1 | C/M/H SOX1 | -RV------- ------Q--- M--E--K--- --------AE --VM--A--- ---D--K--- -L---EH--- -------T- | | AB011802/X94126/Y13436 |
| | C SOX2 | -RV------- ------Q--- M--E--K--- --------AE ------A--- ---D--K--- -L---EH--- -------T- | | J12532 |
| | M/H SOX2 | -RV------- ------Q--- M--E--K--- --------AE ------T--- ---D--K--- -L---EH--- -------T- | | X94127/Z31560 |
| | C SOX3 | -RV------- ------Q--- M--E--K--- --------AD -----DA--- ---D--K--- -V---E---- -------T- | | AB0011803 |
| | M/H SOX3 | -RV------- ------Q--- M-LE--K--- --------AD ----TDA--- ---D--K--- -V---E---- -------T- | | X94125/X71135 |
| | ZF SOX19 | -KV------- ------Q--- M--E--K--- --------AE ----TDA--- ---D--K--- -L---E---- --K----QP | | X79821 |
| B2 | C/M/H SOX14 | ---------- ------Q--- M--E--K--- --------AE ------A--- -Y-D--K--- -----EH--- -------P- | | AB026622/AF193437/AJ006230 |
| | ZF SOX21 | --V------- -----AQ--- M-LD--K--- --------GE -----D---- ---D--K--- -V---EH--- -------P- | | J85092 |
| | C SOX21 | --V------- ------Q--- M--E--K--- --------AE ------A--- ---D--K--- -M---EH--- -------P- | | BAA77266 |
| | H SOX21 | --V------- -----AQ--- M--E--K--- --------AE ----T----- ---D--K--- -M---EH--- -------P- | | AF107044 |
| C | M SOX4 | Q--------- ----QI---- -ME-S----- A--------- ----KD-D-I ---Q------ LK--A----- -----K-V- | | X70298 |
| | H SOX4 | Q--------- ----QI---- -ME-S----- A--------- ----KD-D-I ---R------ LK--A----- -----K-V- | | X70683 |
| | ZF SOX11 | Q--------- ----KI---- -ME-S----- A--------- --M-KD---I ---R------ L---A----- ----KK-P- | | J85090 |
| | C SOX11 | Q--------- ----KI---- -ME-S----- A--------- --M-KD---I ---R------ LK--A----- -----K-P- | | AB102237 |
| | M SOX11 | Q--------- ----KI---- -ME-S----- A--------- --M-KD---I ---R--G--- LK--A----- -----K-P- | | AF009414 |
| | M SOX12 | ----QH---- -MD-W----- A-------R- -Q--QD---I --ER------ LK--A- | | Z18961 |
| | H SOX22 | Q--------- ----QH---- -MD-W----- A-------R- -Q--QD---I --VR------ LK--A----- -----K-S- | | J35612 |
| D | M/H SOX5 | P--------- ---AKD---- -L-AF----- -N---I--S- --AMTNL--Q -YY--QA--S K--LEK---- --K--P-RT | | AJ010604/C2828210 |
| | M SOX6 | P--------- ---AKD---- -L-AF----- -N---I--S- --SM-NQ--Q -YY--QA--S KI-LEK--A- --K--P-RT | | D61689 |
| | M SOX13 | S--------- ---AKD---- -L-AF----- -S---I--S- --SMTNQ--Q -YY--QA--S R--LEK---- --K--P-RT | | AJ000740 |
| E | C SOX8 | P-V------- ---AQAA--- L-D-Y-HL-- A-L--T---L -R----N--- --V------- V--K--H--- --Q---R-S | | AF228664 |
| | M/H SOX8 | P-V------- ---AQAA--- L-D-Y-HL-- A-L--T---L -R-------- --V------- V--K--H--- --Q---R-S | | AF191325/HS249H11 |
| | C SOX9 | P-V------- ---AQAA--- L-D-Y-HL-- A-L--T---L -R--N----- --V------- V--K--H--- --Q---R-S | | P48434 |
| | M/H SOX9 | P-V------- ---AQAA--- L-D-Y-HL-- A-L--T---L -R--N----- --V------- V--K--H--- --Q---R-S | | Q04887/Z46629 |
| | C SOX10 | P-V------- ---AQAA--- L-D-Y-HL-- A-L--T---L -R--N--D-- ---------- M--K--H--- --Q---W-N | | AAD38050 |
| | M/H SOX10 | P-V------- ---AQAA--- L-D-Y-HL-- A-L--T---L -R--N--D-- ---------- M--K--H--- --Q---R-N | | AF047043/AF001183 |
| F | M SOX7 | SR-R------ ---AKD--KR L-V----L-- A-L--M---S --A-TL-Q-- -YVD------ L---Q---A- --------Q | | AB023419 |
| | M SOX17 | SR-R------ ---AKD--KR L------L-- A-L--M---S --A-TLA--- --V------- V---Q-H-A- ------R-Q | | D49474 |
| | M SOX18 | LR-R------ ---AKD--KR L------L-- AVL--M---A --C-NTA--- --V------- V--LR-H-A- --------Q | | L35032 |

```
G       M SOX15       EKV------- ----SVQ--Q M-----E--- --------AQ ----GDE--- --V---K--- -R-LR----- -------S-   AF182945
        M SOX16                  ----SAQ--Q M-----K--- --------AQ ----DDE--- --V---K--- -R-LH--               L29084
        H SOX20       EKV------- ----SAQ--Q M-----K--- --------AQ ----D-D--- --V---K--- -R-LR----- -------A-   AB006867

H       H SOX30       Q-V------- ---A-IH-PA L-KA--AAA- A---VQ--LE -NK---EQ-K -YYD--QKIK EK-REEF-GW V-Q--PG-R   AB022083
I       X SOX31       QLV------- ----S---KR MSALH-K--- ----R---EI -RG-G-EDR- --R---K--- ---AI-F-G- --A--K-R-   BAA32249
J       Ce SOXJ       PRI------- ----QQR-QQ I-ATGQ-F-- -D---M--AE -RKME-H--V --V-R-KQ-- EE-FNAH--- V-----R-R   J51998

Outgroups

        M LEF1        P---K-L--- -LYMKEM-AN VVAECTLKES AA-NQI--R- -HA--RE-QA KYY-L-RKE- QL--QL--GW SA-DNYG--   P27782
        M TCF1        PV--K-L--- -LYMKEM-A- VIAECTLKES AA-NQI--R- -HA--RE-QA KYY-L-RKE- QL--QL--GW SA-DNYG--   Q00417
```

**Figure 1.1:** Alignment of HMG domain sequences, modified from (Bowles et al., 2000)

HMG domains of Sox protein representatives for each subgroup from zebrafish, *Danio rerio* (ZF), chicken, *Gallus gallus* (C), mouse, *Mus musculus* (M), human, *Homo sapiens* (H), *Xenopus laevis* (X) and *Caenorhabditis elegans* (Ce) were aligned with ClustalW. The subgroups are labelled A-J and the accession numbers for each gene are shown (Acc. No.). Identical amino acids to the consensus sequence (top) are represented by dashes. Amino acids conserved across all Sox genes are underlined in the consensus sequence, whereas residues which form the 3 α-helices are bracketed. Mouse LEF1 and TCF1 sequences are included as outgroups.

that all 3 C-terminal domains by themselves show strong transactivation of Chloramphenicol Acetyl Transferase (CAT) regardless of cell-types when fused to the GAL4 DNA binding domain (Ng et al., 1997; Kamachi et al., 1999; Kamachi et al., 2000). The N-terminal synergy domain of some Sox proteins such as Sox10 was shown to be required for interaction with partner factors (Kuhlbrodt et al., 1998a), but in addition, the region just C-terminal to the HMG domain also makes a minor contribution (Kamachi et al., 1999). Partner transcription factors might then confer tissue specificity of transcriptional regulation by their restricted expression. Furthermore, specificity might be improved by stabilising the binding of the SOX protein to its DNA target site by the partner factor. Such binding sites for SOX proteins and partner factors have been identified on the δ-crystallin enhancer for SOX2 and δEF3 (Kamachi et al., 1995; Kamachi et al., 1999); SOX2 and Oct3/4 on the FGF4 enhancer (Dailey et al., 1994; Ambrosetti et al., 1997); SOX9, SF1, GATA and WT1 on the anti-Muellerian hormone (AMH) enhancer (Marshall and Harley, 2000); SOX9, L-SOX5 and SOX6 on the Col2a1 minimal enhancer (Lefebvre et al., 1998); and SOX10 and PAX3 on the MITF promoter (Bondurand et al., 2000; Potterf et al., 2000).

## *Aims of the project*

The neural crest has been recognised as a highly important embryonic structure that gives rise to a great number of various derivatives such as neurons and glia of the peripheral nervous system (PNS), the enteric nervous system (ENS), craniofacial cartilage structures and pigment cells. The underlying processes and timing involved in specifying those different cell fates from multipotent precursors is still largely unknown. Not surprisingly, there are numerous human congenital diseases, generally termed neurocristopathies, in which either specification, survival, proliferation or differentiation of one or more neural crest derivatives has been disrupted.
Waardenburg Shah Syndrome (WS4) patients and their mouse model, the *Dominant megacolon* mouse (*Dom*), present with defects in most neural crest derivatives and are caused by lesions in the transcription factor *SOX10* (Herbarth et al., 1998; Pingault et al., 1998a; Southard-Smith et al., 1998). As a result, SOX10 was suggested to be a key factor in neural crest development (Pevny and LovellBadge, 1997).
The phenotype of the zebrafish pigment mutant *colourless* (*cls*; Kelsh et al., 1996; Malicki et al., 1996) is highly reminiscent of these mammalian *Sox10* mutants

(described in detail in a previous section of this chapter) and thus, *sox10* was proposed to be a promising candidate gene for the *cls* locus.

The aim of this project is to investigate whether *cls* encodes a zebrafish *sox10* orthologue. This was attempted by

i) Cloning a *sox10* homologue from zebrafish embryos

ii) Testing whether *sox10* maps to the same locus on LG3 as *cls*

iii) Testing for mutant lesions in at least one of the *sox10* alleles

iv) Attempting to rescue the colourless mutant phenotype by injecting wild-type *sox10* cDNA inducible by a heatshock promoter

v) Analysis of the embryonic expression pattern of wild-type and *cls* mutants knowing that *cls* acts cell-autonomously at least within the pigment lineage (Kelsh and Eisen, 2000b).

# Chapter 2 – Materials and Methods

## *Fish husbandry*

Wild-type and mutant embryos were obtained from the University of Oregon. Crosses were set up as described in Westerfield (1995), eggs were collected from successful matings on the following morning with the use of a tea-strainer. After rinsing with RO water, eggs were transferred to a Petri-dish containing embryo medium (5mM NaCl, 0.17mM KCl, 0.33mM $CaCl_2$, 0.33mM $MgSO_4$, $10^{-5}$% methylene blue to suppress fungal growth). The embryos were raised at 28.5°C and staged according to Kimmel et al. (1995). Embryos to be live mounted or sorted were anaesthetised in a 0.2% solution of 3-aminobenzoic acid ethyl ester ("Tricaine", Sigma) prior to analysis.

## *Primer design and sequences*

Primers were designed with the use of two programs. First, suitable primer pairs for PCR amplifications were suggested by Primer3 (http://waldo.wi.mit.edu/cgi-bin/primer/primer3.cgi/primer3_www.cgi).
Primer pairs, each containing at least 1-2 GC clamps at the 3'end were then analysed for hairpins, dimers, crossdimers and melting temperature (Tm) with a second program, NetPrimer (http://www.premierbiosoft.com/netprimer/netprimer.html).
Only primers with a rating of at least 80 (out of 100; NetPrimer) were selected.
Sequencing primers were directly designed and analysed with NetPrimer. Primers S21, S22 and S23 were obtained from MWG-Biotech AG, 10nmol scale of synthesis and HPSF purification. All other primers were ordered from Gibco BRL Lifetechnologies with standard purification and 50nmol synthesis scale. Primers used for site-directed mutagenesis were additionally PAGE purified.

### Primers used to clone and map *sox10*

5'SOX 10*: 5' AAG GCC GGA TCC ATG AAY GCN TTY ATG GTN TGG 3'
3'SOX 10*: 5' AAG GCC GGA TCC GGY TGR TAY TTR TAR TCN GG 3'

5GSP: 5' TCC GTC TCG TTC AGC AGT CTC CAC AGC 3'

N5GSP: 5' GGC GTT GTG CAG GTG CGG ATA TTG ATC C 3'

3GSP: 5' AAG CTG TGG AGA CTG CTG AAC GAG ACG G 3'

N3GSP: 5' TAT CGA GGA GGC CGA GCG CTT GAG G 3'

S9: 5' AGG TCT GGC TGG AGT GGG CGT AGT AGG 3'

S10: 5' CGA TGA TTT TTA GCA CAC ACA CAC ACC 3'

S11: 5' ACC GTG ACA CAC TCT ACC AAG ATG ACC 3'

S12: 5' AAA ATC ATC GAA ACA CTC GCC TGC ACC 3'

S13: 5' CAT GAT AAA ATT TGC ACC CTG AAA AGG 3'

S14: 5' GTA TTT ATT TAC TTA CCC AAT GTT AGG 3'

*...In those degenerate primers, the letter "Y" stands for C or T, "R" for A or G, "N" could be either A, C, G or T.

## Primers used to search for mutant lesions

S19: 5' GCA GCA AGA GCA AAC CGC ACG 3'

S20: 5' TGG TAG GGG GCG TTG GAG GGC 3'

S21: 5' ACC TAC CGA AGT CAC CTG TGG 3'

S22: 5' GAT ATT GAT CCG CCA GTT TCC 3'

S24: 5' AAT CGC ATT ACA AGA GCC TGC 3'

S25: 5' CCA GGG AAG TGT GTT TCA CTC 3'

S26: 5' TAT ACA TAC GGC ATC TCC AGC 3'

S27: 5' AGT TTG TGT CGA TTG TGG TGC 3'

## Primers used to make the *sox10* heatshock construct

*Cla* I-S21: 5' CCA TCG ATA CCT ACC GAA GTC ACC TGT GG 3'

S27-*Xba* I: 5' GCT CTA GAG TTT GTG TCG ATT GTG GTG C 3'

## Primers used to make the *sox9a* and *sox9b* heatshock constructs

S9A-1: 5' GGA TCC AAG CTT ATC GAT TTC G 3'

S9A-2: 5' GCT CTA GAG CTT TTT CAG TGC ACA TTC AGG C 3'

S9B-1: 5' CCA TCG ATG GAT CTG TGT GTG TTT CAG CAG C 3'

S9B-2: 5' GCT CTA GAG CTG CTG ATA GTG ATT CAG GCG 3'

### Primers used to sequence the *sox9a* and *sox9b* constructs

9A/B19: 5' GTG CTG AAG GGY TAC GAC TGG 3'

9A20: 5' GGT GGT GGG AGG AGT GGG CGG 3'

9A22: 5' CCT GCG CCC ACA CCA TAA ACG 3'

9A24: 5' ACC CAC ATC TCG CCC AAC GCC 3'

9A25: 5' GCT CGT TGT AAT GGC TCG GGC 3'

9A26: 5' CAA AAC GGC AGC CCT CAA AGC 3'

9B20: 5' GGT GGT GGG TGG TGT GGG GGG 3'

9B22: 5' CTT GAG CCC AAA CCA TAA ACG 3'

9B24: 5' ATG CGC TGT TCA GAG CCC TGC 3'

9B25: 5' CCT CTG CTG GTG CTG CTC GCC 3'

9B26: 5' AGT ACC TGC CTC CGC ACG GGG 3'

### Primers used for site-directed mutagenesis

m618→WT1: 5' GCT CAG CAA AAC ACT GGG GAA GCT GTG GAG ACT GC 3'

m618→WT2: 5' GCA GTC TCC ACA GCT TCC CCA GTG TTT TGC TGA GC 3'

WT1→tw2: 5' GGGCGC AGA TGG CGG GTA AAC GCA GAT AAA GAG 3'

WT2→tw2: 5' CTC TTT ATC TGC GTT TAC CCG CCA TCT GCG CCC 3'

WT1→Y83X: 5' AGG TGC TGA ACG GGT AAG ACT GGA CGC TCG TGC 3'

WT2→Y83X: 5' GCA CGA GCG TCC AGT CTT ACC CGT TCA GCA CCT 3'

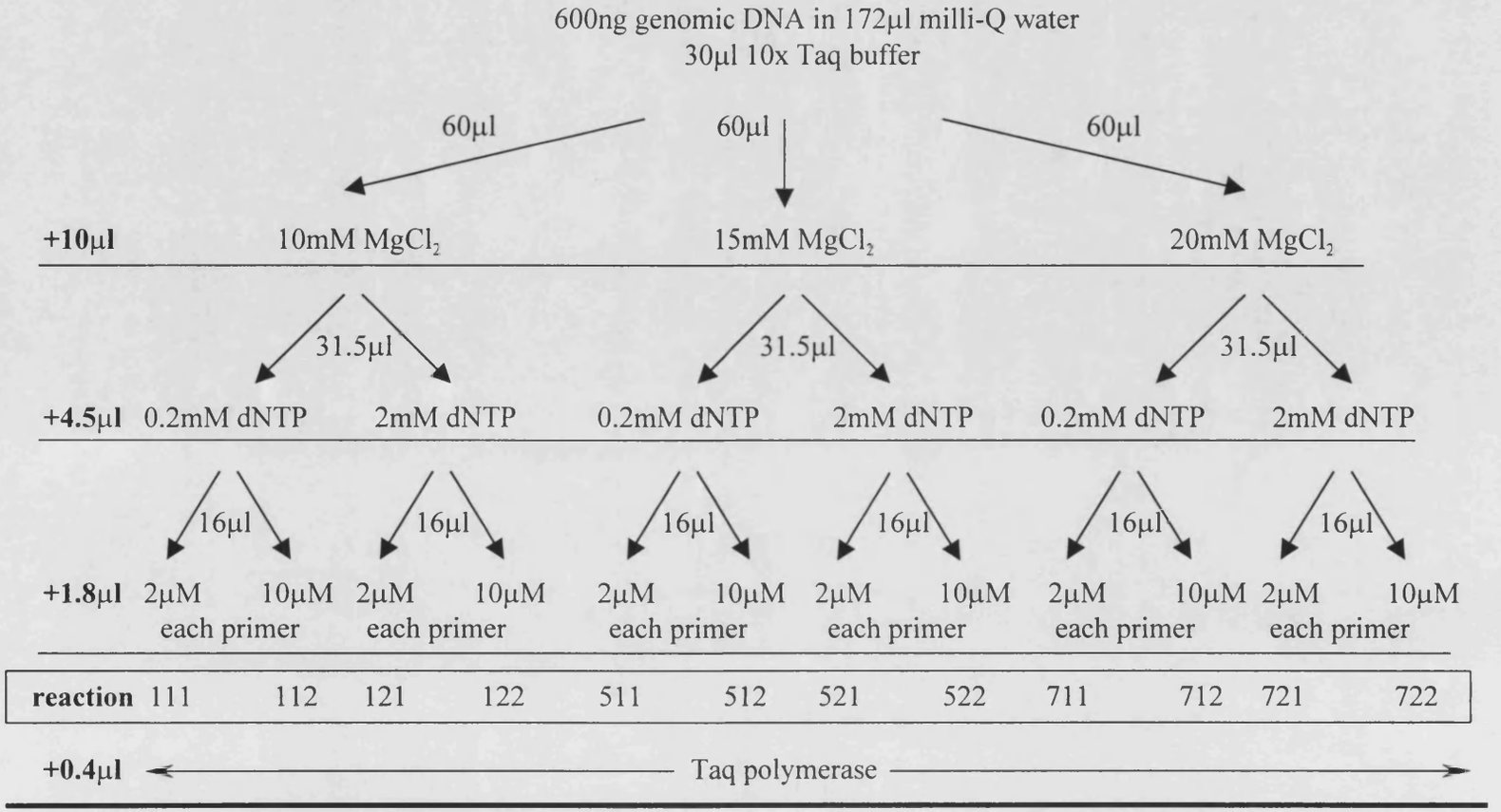
WT1→E189X: 5' AGC TCA GAG GCC TAG GCC CAC TCT GAG GGT GAG G 3'

WT2→E189X:5' CCT CAC CCT CAG AGT GGG CCT AGG CCT CTG AGC T 3'

## *Polymerase Chain Reactions (PCR)*

PCR conditions ($MgCl_2$, dNTP and primer concentrations) were optimised with a scheme of 12 reactions varying all parameters in combinations (Figure 2.1). The Tm calculated by NetPrimer, determined by nearest neighbour analysis, was used as the annealing temperature.

# PCR Optimisation scheme




12 x 20µl reactions

**Figure 2.1:** This PCR optimisation scheme (designed by Stone Elworthy, University of Bath) varies 3 concentrations of magnesium chloride in combination with 2 concentrations each of dNTPs and primers. The molarities noted in this scheme represent the concentrations of stock solutions of $MgCl_2$, dNTP and primers. The resulting final test conditions used in each of the 12 reactions were as follows:

| | **111** | **112** | **211** | **212** | **511** | **512** | **521** | **522** | **711** | **712** | **721** | **722** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **$MgCl_2$** | 1mM | 1mM | 1mM | 1mM | 1.5mM | 1.5mM | 1.5mM | 1.5mM | 2mM | 2mM | 2mM | 2mM |
| **dNTPs** | 20µM | 20µM | 200µM | 200µM | 20µM | 20µM | 200µM | 200µM | 20µM | 20µM | 200µM | 200µM |
| **primers** | 0.2µM | 1µM | 0.2µM | 1µM | 0.2µM | 1µM | 0.2µM | 1µM | 0.2µM | 1µM | 0.2µM | 1µM |

## RT-PCR Protocols

### *Isolation of total RNA using Tri Reagent*

Approximately 200-300 embryos of the desired stage were dechorionated using watchmaker forceps, killed by an overdose of the anaesthetic Tricaine and homogenised thoroughly in 1ml Tri Reagent (Sigma). Homogenates were spun at 13000rpm for 10 minutes at 4°C, the supernatant was removed to a fresh tube and the sample allowed to stand for 5 minutes at room temperature. 0.2ml chloroform was added, the tube shaken vigorously and left to stand for 2-15 minutes at room temperature. After centrifugation at 13000rpm for 15 minutes at 4°C the aqueous phase containing RNA was transferred to a fresh tube, 1/10 volume of isopropanol was added, the tube was inverted a few times and allowed to stand at room temperature for 5 minutes. The sample was spun again at 13000rpm for 10 minutes at 4°C and an equal volume of isopropanol added to the transferred supernatant. After the tube was inverted again and left to stand for 5-10 minutes at room temperature the precipitated total RNA was pelleted at 13000rpm for 10 minutes at 4°C. The supernatant was removed, the pellet washed with 1ml 75% ethanol by vortexing and the sample respun at 13000rpm for 5 minutes at 4°C. After removal of all the ethanol, the pellet was dried at room temperature for approximately 5 minutes and redissolved in a suitable volume (approximately 50µl) of fresh milliQ water. The quality of the total RNA preparation was assessed by analysing 1µl of the sample on a standard 1%TBE gel.

### *Reverse transcription using random hexamers or an oligo dT primer*

3-6µg of total RNA was combined with 100ng random hexamers or 500ng oligo dT and the volume made up to 12µl with fresh milliQ water. The sample was heated to 70°C for 10 minutes, chilled on ice, then 4µl 5x reaction buffer (Gibco BRL), 2µl 0.1M DTT and 1µl 10mM dNTPs were added. After mixing the reaction, samples were incubated at 25°C (or 42°C for oligo dT reactions) for 2 minutes and 1 µl (200u) Superscript II (Gibco BRL) was added. When using random hexamers, tubes were incubated for another 10 minutes at 25°C. All reactions were transferred to 42°C for 50 minutes followed by 15 minutes at 70°C to inactivate the reverse transcriptase. Then 2u RNAse H (Promega) was added and reactions incubated at 37°C for 20 minutes. The reactions

were diluted to a 1:50 final concentration directly into PCR reactions (i.e. 1µl cDNA used in a 50µl PCR reaction).

## PCR protocols for primer combinations used to clone *sox10*

### *Degenerate primers 5'Sox10 and 3'Sox10 to clone the HMG box*

These primers and the PCR protocol were first published by Yuan et al. (1995). A 50µl PCR reaction included 4µl 25mM (2mM final) $MgCl_2$, 5µl 2.5mM (250µM final) dNTP, 5µl 10µM (1µM final) of each primer, 1µl (5u) Taq DNA polymerase (Promega), 5µl 10x reaction buffer and 2µl cDNA. Reactions were subjected to 1 minute at 94°C, 25 cycles of 1 minute at 94°C, 1 minute at 45°C and 1 minute of 72°C followed by a final extension step of 10 minutes at 72°C.

### *RACE-PCR protocol*

RACE-PCR (Rapid Amplification of cDNA Ends-Polymerase Chain Reaction) was carried out using the Clontech SMART (Switching Mechanism At 5' end of RNA Transcript) RACE cDNA Amplification Kit. First, 5'RACE cDNA and 3'RACE cDNA were synthesised by combining 1µg of total RNA from 19hpf embryos with 1µl 5'-CDS primer or 3'-CDS primer respectively and 1µl SMART II (5'RACE cDNA reaction only). Supplied sterile water was added to the final volume of 5µl, the reactions incubated at 70°C for 2 minutes and cooled on ice for 2 minutes. After spinning the tubes briefly to collect the contents, 2µl 5x First-Strand buffer, 1µl 20mM DTT, 1µl 10mM dNTP mix and 1µl (200u) Superscript II (Gibco BRL) reverse transcriptase were added. Reactions were incubated at 42°C for 90 minutes, diluted with 100µl Tricine-EDTA buffer supplied with the kit and heat treated at 72°C for 7 minutes.
RACE PCR reactions were set up by combining 2.5µl 5'RACE or 3'RACE cDNA with 5µl 10x UPM (universal primer mix), 1µl 10µM 5GSP (5'gene specific primer, for sequence see primer design and sequence section of Materials and Methods) or 3GSP (3' gene specific primer), 5µl 10x Reaction Buffer, 1µl 10mM dNTP mix, 1µl 50x Advantage 2 Polymerase Mix (Clontech) and 34.5µl PCR-grade water. In a MJ Research PTC-DNA Engine PCR reactions were subjected to 5 cycles of 94°C for 10 seconds, 72°C for 3 minutes, then 5 cycles of 94°C for 10 seconds, 70°C for 10 seconds,

72°C for 3 minutes, then 30 cycles of 94°C for 10 seconds, 68°C for 10 seconds, 72°C for 3 minutes followed by a final extension step at 72°C for 10 minutes. 5µl of each reaction were run out on a 1% agarose/ethidium bromide gel and examined using an ultraviolet light source.

To test for product specificity nested PCR reactions were carried out by diluting 5µl of the primary PCR products into 245µl of Tricine-EDTA buffer. PCR reactions were set up as before, but replacing the 2.5µl cDNA with 5µl of the diluted primary PCR product and 1µl 10µM 5GSP or 3GSP with either 1µl 10µM N5GSP or N3GSP, respectively. For primer sequences see primer design and sequence section of Materials and Methods. The nested PCR reactions were subjected to 25 cycles of 94°C for 10 seconds, 68°C for 10 seconds and 72°C for 3 minutes followed by a final extension of 72°C for 10 minutes. As before, 5µl of each reaction were run out on a 1% agarose/ethidium bromide gel (TBE) and examined on an ultraviolet light source.

### *PCR protocol for amplification between primers 5'Sox10, 3GSP or N3GSP and S13 (clone 15 and 20)*

Since the 3'RACE product did not span the entire region 3' to the HMG box, an overlapping fragment was amplified between primers 5'Sox10 and S13. A 20µl PCR reaction included 2µl cDNA, 2µl 10µM (1µM final) of each primer, 0.3µl (1.5u) Thermoprime Taq DNA polymerase (Advanced Biotechnologies), 1.8µl 11x reaction buffer [167µl 2M TrisHCl pH8.8, 83µl 1M Ammonium sulphate, 33.5µl 1M Magnesium chloride, 3.4µl 10mM EDTA pH8, 3.6µl β-mercaptoethanol, 75µl of each 100mM dNTP stock and 85µl 10mg/ml BSA (DNAse free, Pharmacia), recipe kindly provided by Dr. Diane Hird, University of Bath] and 11.9µl milli-Q water to make up to 20µl. Reactions were subjected to 1 minute at 94°C, 35 cycles of 30 seconds at 94°C, 1 minute at 60°C and 3 minutes 30 seconds of 72°C followed by a final extension step of 10 minutes at 72°C.

Subsequent nested PCRs involved amplification between primers 3GSP or N3GSP and S13. The PCR reaction were set up like described in the previous section for the 5'Sox10/S13 amplification replacing the 2µl cDNA with equal volume of 1:10 diluted first round PCR template. Reactions were subjected to 1 minute at 94°C, 25 cycles of 30 seconds at 94°C, 1 minute at 66°C and 3 minutes 30 seconds of 72°C followed by a

final extension step of 10 minutes at 72°C. Products obtained from this amplification were cloned into pGEM-T Easy vector and sequenced.

## PCR protocols involved in mapping *sox10* on the radiation hybrid panel LN54

The LN54 collection of radiation hybrids (zebrafish in a mouse background) was kindly provided by Marc Ekker from the Loeb Research Institute in Ottawa, Canada (Hukriede et al., 1999). It contains 93 radiation hybrid DNA samples (100ng/µl each), DNA from the two parental cell lines (zebrafish AB9 and mouse B78) and a 1:10 mixture of the latter two.

For mapping *sox10* on the panel, a PCR reaction with primers S11 and S13 was used. For primer sequences see primer design and sequence section of Materials and Methods. 1µl of each of the 96 DNA samples was combined with 2µl 10µM S11 primer, 2µl 10µM S13 primer, 1.8µl 11x PCR buffer (recipe see previous section), 0.3µl Thermoprime Taq DNA polymerase (Advanced Biotechnologies) and 12.9µl milliQ water.

PCR reactions were denatured at 94°C for 1 minute, then cycled at 94°C for 30 seconds, 66°C for 1 minute, 72°C for 3 minutes for 30 cycles, followed by a final extension of 72°C for 10 minutes. To increase the intensity and specificity of the obtained products, those first PCR products were diluted 1:10 and 1µl was then included in a nested PCR reaction. The reaction set up and PCR program was identical to the one outlined above with the exception that 2µl 10µM primer S11 were replaced by 2µl 10µM of the nested primer S12. All 20µl of nested PCR reactions were loaded on a 1% agarose/ethidium bromide gel and location of the products of the expected 850bp size was recorded. This assay was carried out in duplicate and the resulting data submitted to the mapping program at http://mgchd1.nichd.nih.gov:8000/zfrh/beta.cgi.

## PCR protocols for primer pairs used to screen for mutant lesions

The sequence of mutant *cls* alleles was compared to wild-type *sox10* by sequencing 4 overlapping PCR fragments spanning the entire coding region. Those fragments were amplified between primer pairs S19-S20, S21-S22, S24-S25 and S26-S27. The optimum

conditions were determined by setting up a PCR optimisation scheme (Figure 2.1) for each primer pair. S19-S20 and S21-22 gave cleanest products with condition "122", the other primer pairs worked optimally in condition "121" (see Figure legend 2.1).
A 50µl "122" reaction contained 5µl mutant or wild-type cDNA, 5µl 10x reaction buffer, 1µl 50mM (1mM final) $MgCl_2$, 1µl 10mM (200µM final) dNTP, 5µl 10µM (1µM final) of each primer, 1µl (5u) Taq DNA polymerase (GibcoBRL) and milli-Q water to make up to 50µl.
A 50µl "121" reaction contained 5µl mutant or wild-type cDNA, 5µl 10x reaction buffer, 1µl 50mM (1mM final) $MgCl_2$, 1µl 10mM (200µM final) dNTP, 1µl 10µM (0.2µM final) of each primer, 1µl (5u) Taq DNA polymerase (GibcoBRL) and milli-Q water to make up to 50µl.
The PCR program identical for all primer pairs apart from the annealing temperature involved 1 minute at 94°C, 35 cycles of 30 seconds at 94°C, 1 minute at 68°C for S19-S20, 58°C for S21-S22, 56°C for S24-S25 or 55°C for S26-S27 and 1 minute at 72°C followed by a final extension step of 10 minutes at 72°C.

## PCR protocol for primers used to make the *sox10* heatshock construct

*sox10* coding regions were amplified from wild-type and *cls* mutant cDNAs using primers *Cla* I-S21 and S27-*Xba* I.
In a 20µl reaction, 0.4µl undiluted wild-type cDNA from 38hpf total RNA was combined with 0.4µl (2u) Taq DNA polymerase (TaqPlus Precision PCR System, Stratagene), 2µl 10x reaction buffer, 2µl 2mM (200µM final) dNTPs, 2µl of each 2µM (0.2µM final) primer and 11.2µl milli-Q water.
Mutant coding regions of alleles $cls^{m618}$ and $cls^{tw2}$ were amplified in 10µl reactions containing 1µl 1:5 diluted mutant cDNA, 0.2µl 10mM (200µM final) dNTPs, 1µl 2µM (0.2µM final) of each primer, 0.2µl (1u) Herculase (Herculase enhanced DNA Polymerase, Stratagene), 1µl 10x reaction buffer and 5.6µl milli-Q water.
All PCR reactions were denatured at 92°C for 1 minute, then subjected to 10 cycles of 92°C for 10 seconds, 67°C to 57°C (1°C/cycle) for 45 seconds, 72°C for 2 minutes

followed by 30 cycles of 92°C for 10 seconds, 57°C for 45 seconds and 72°C for 2 minutes followed by a final extension of 72°C for 10 minutes.

## PCR protocol for primers used to make the *sox9a* and *sox9b* heatshock constructs

*sox9a* and *sox9b* coding regions were amplified using primers S9A-1 and S9A-2 or S9B-1 and S9B-2, respectively.

10μl reactions were set up with 200pg of each plasmid miniprep containing the full length *sox9a* and *sox9b* clones, 0.2μl 10mM (200μM final) dNTPs, 1μl 10μM (1μM final) of each primer, 0.2μl (1u) Herculase (Herculase enhanced DNA Polymerase, Stratagene), 1μl 10x reaction buffer and 5.6μl milli-Q water.

PCR reactions amplifying the *sox9a* coding region were denatured at 92°C for 1 minute, then subjected to 10 cycles of 92°C for 10 seconds, 68°C to 58°C (1°C/cycle) for 45 seconds, 72°C for 2 minutes followed by 25 cycles of 92°C for 10 seconds, 58°C for 45 seconds and 72°C for 2 minutes followed by a final extension of 72°C for 10 minutes.

PCR reactions amplifying the *sox9b* coding region were denatured at 92°C for 1 minute, then subjected to 10 cycles of 92°C for 10 seconds, 60°C to 48°C (1.2°C/cycle) for 45 seconds, 72°C for 2 minutes 30 seconds followed by 25 cycles of 92°C for 10 seconds, 48°C for 45 seconds and 72°C for 2 minutes and 30 seconds followed by a final extension of 72°C for 10 minutes.

## Protocol for primers used for site-directed mutagenesis

To create a wild-type *sox10* heatshock construct from the $cls^{m618}$ mutant heatshock construct, primers m618→WT1 (forward primer) and m618→WT2 (reverse primer) were used. The $cls^{tw2}$ mutant heatshock construct was then created from the wild-type construct using primers WT1→tw2 (forward primer) and WT2→tw2 (reverse primer).

Four 50μl reactions were set up, each containing either 5ng, 10ng, 20ng or 50ng of template ($cls^{m618}$ mutant heatshock construct to make the wild-type construct and wild-type to make the $cls^{tw2}$ mutant heatshock construct), 1μl 125ng/μl of each primer, m618→WT1 and m618→WT2 or WT1→tw2 and WT2→tw2, 1μl 20mM (400μM final) dNTPs, 1μl (2.5u) Pfu Turbo DNA Polymerase (Stratagene), 5μl 10x reaction

buffer and milli-Q water to make the reaction volume up to 50µl. A "no primer" control reaction was included containing 20ng of template, but no primers.
Reactions were subjected to 95°C for 30 seconds and 12 cycles of 95°C for 30 seconds, 55°C for 1 minute and 68°C for 12 minutes and 30 seconds.
10µl of each reaction were analysed on a standard agarose gel for successful amplification. To the remaining 40µl reaction, 1µl Dpn I restriction enzyme (NEB) was added directly into the PCR tubes and samples incubated at 37°C for 60 minutes to digest the parental methylated strands. 1.5µl of those double-stranded nicked open circular plasmids containing the mutated site were transformed into Epicurian Gold Supercompetent cells (Stratagene) and colonies tested for successful mutagenesis by either restriction digest with *Tsp*RI or by sequencing across the mutated region. If the wild-type construct was created successfully, a *Tsp*RI restriction was gained compared to the $cls^{m618}$ mutant allele. To test this, a 20µl restriction digest with 2.5µl putative wild-type construct miniprep, 0.2µl BSA, 1µl *Tsp*RI (NEB), 2µl 10x NEB4 reaction buffer and 14.3µl milli-Q water was set up and incubated at 65°C for 3 hours. Positive clones were sequenced to ensure successful mutagenesis.
Since there was no similar test based on a change of restriction digest pattern available for the $cls^{tw2}$ heatshock construct, 350ng of several minipreps were directly sequenced across the mutagenesis site.

## *Molecular Cloning Techniques*

### Agarose gel electrophoresis

To prepare a 1% agarose gel, 0.5g agarose was weighed into 50ml 1xTBE buffer [per litre 10.8g Tris base, 5.5g boric acid, 4ml 0.5M EDTA pH8.0] and heated in a microwave at full power for 2 minutes to dissolve the agarose. The evaporated liquid was replenished with RO water, the gel solution briefly cooled under tap water, 3µl (10mg/ml) ethidium bromide added and the gel poured. Gels were run at 80-100V for approximately 1 hour in 1x TBE buffer with additional 5µl ethidium bromide added to the cationic compartment of the gel tank, examined using an ultraviolet light source and documented using a black and white video camera (UVP) and a video graphic thermal printer (Sony).

## Restriction digests

Approximately 1μg of plasmid DNA or 200-300ng of a PCR product were digested in a total reaction volume of usually 10μl containing 1μl 10x restriction enzyme buffer and 10 units of restriction enzyme. Incubation was carried out at the enzyme's specific reaction temperature (usually 37°C) for at least 2 hours and up to overnight. The majority of enzymes were purchased from Promega or New England Biolabs (NEB).

## Ligation of PCR products

50-100ng of PCR product was directly ligated into 25ng of pGEM-T Easy vector (Promega) in a total reaction volume of 20μl containing 3units T4 DNA ligase (Promega) and supplied ligation buffer. The reaction was carried out at 16°C at least overnight or preferably for the duration of 2 days.

## Preparation of $CaCl_2$ competent cells

DH5α cells from a glycerol stock were plated on a fresh LB agar plate [LB medium contains 10g bacto-tryptone, 5g bacto-yeast extract, 10g NaCl per litre with pH adjusted to 7.0 with 5N NaOH; for LB-agar plates 15g bacto-agar is added] and incubated at 37°C overnight. 50ml LB medium was then inoculated with one colony and again incubated at 37°C in a shaking incubator overnight. 2ml of overnight culture were diluted into 50ml fresh LB medium and incubated at 37°C in a shaking incubator until $A_{600}$ = 0.4-0.6. The culture was pelleted at 4000rpm for 10 minutes at 4°C and the supernatant discarded. The well drained cell pellet was resuspended in 20ml ice cold 50mM $CaCl_2$ and left on ice for at least 20 minutes. The cells were centrifuged again at 4000rpm for 10 minutes at 4°C, the supernatant discarded and the cells resuspended in a solution of 4.25ml ice cold 50mM $CaCl_2$ and 0.75ml sterile glycerol. Aliquots of 200μl were snap-frozen in a dry ice/acetone bath and stored at –80°C.

## Transformation of competent *E.coli* by heatshock

### *Transformation of competent DH5α cells prepared with the $CaCl_2$ method*

A 200µl aliquot of competent DH5α cells was thawed on ice, 20µl ligation reaction was added and gently mixed with a pipette tip. The tube was kept on ice for 30 minutes followed by heatshock in a waterbath at 42°C for exactly 90 seconds and returned to ice for another 2 minutes. 800µl LB medium was added and the cells left to recover at 37°C for an hour. The cells were pelleted at 5000rpm for 3 minutes, resuspended in approximately 50µl LB medium and spread onto an LB agar plate containing 50µg/ml ampicillin (Sigma), 20µg/ml X-gal and 20µg/ml IPTG.

### *Transformation of Gold Super-competent Epicurian Coli cells*

Constructs created by site-directed mutagenesis were transformed into Gold Super-competent Epicurian Coli cells (Stratagene) according to the manufacturers instructions. Cells were slowly thawed on ice, gently swirled to mix cells and 1µl XL10-Gold β-mercaptoethanol was added. During the following 10 minute incubation on ice the cells were gently swirled every 2 minutes and then 1.5-2.0µl ligation reaction was added. After incubating the cells on ice for 30 minutes, they were heatshocked at 54°C for 60 seconds, returned to ice for 2 minutes and 150µl LB medium was added. The cells were allowed to recover at 37°C for an hour and were then spread onto an LB agar plate containing 50µg/ml ampicillin (Sigma).

## Plasmid preparations

### *Plasmid minipreps using alkaline lysis*

The protocol for Lysis by Alkali described in Sambrook et al. (1989) was followed with a few alterations. Single colonies were picked from transformation plates and grown overnight in a 37°C shaking incubator in 2.5ml LB medium containing 50µg/ml ampicillin (Sigma). The culture was centrifuged at 13000rpm for 3 minutes, the pellet drained well, resuspended in 100µl resuspension buffer [5mM glucose, 10mM EDTA pH 8.0 and 100mM Tris/HCl pH 8.0] by vortexing and left to stand at room temperature for 10 minutes. Then 200µl of freshly prepared lysis mix [0.2M NaOH, 1% SDS] was

added and the tubes sharply inverted until the solution was clear and viscous. After 5 minutes incubation on ice, 150µl neutralisation solution [3M NaOAc pH 4.8] was added, the tubes inverted as before and stored on ice again for 5 minutes. Samples were centrifuged at 13000rpm for 10 minutes, the clear supernatant was transferred to a fresh Eppendorf tube, the plasmid DNA was precipitated with 1ml 100% ethanol for at least 20 minutes at –20°C and collected by centrifuging at 13000rpm for 10 minutes. The pellet was resuspended in 100µl 0.1M NaOAc pH 4.8 and reprecipitated for at least 20 minutes at –20°C by adding 200µl 100% ethanol. The DNA was pelleted again by spinning at 13000rpm for 10 minutes, the ethanol removed and finally, the dried DNA pellet was taken up in a solution of 36µl milliQ water and RNAse (4µl 1mg/ml). The sample was incubated at 37°C for 30 minutes and then stored at –20°C until needed.

### *Plasmid minipreps using the Wizard Plus SV Miniprep DNA Purification System (Promega)*

Single colonies were picked from transformation plates and grown overnight in a 37°C shaking incubator in 5ml LB medium containing 50µg/ml ampicillin (Sigma). The culture was centrifuged at 13000rpm for 5 minutes, the pellet drained well and completely resuspended in 250µl Wizard Plus Cell Resuspension Solution by vortexing thoroughly. 250µl Wizard Plus Cell Lysis Solution was added and mixed by inverting the tube until lysis was complete. 10µl Alkaline Protease Solution was added, the tube mixed by inverting it 4 times and incubated at room temperature for 5 minutes. Then 350µl Wizard Plus Neutralisation Solution was added, the tube inverted 4 times and spun at 13000rpm for 10 minutes at room temperature. The cleared lysate was transferred into a Wizard Spin Column and centrifuged at 13000rpm for 1 minute at room temperature. The flow through was discarded, 750µl Wizard Plus Column Wash Solution was added to the column and respun at 13000rpm for 1 minute at room temperature. The last step was repeated with 250µl Wash Solution and the column spun at 13000rpm for 2 minutes at room temperature. The column was then transferred to a fresh Eppendorf tube, the plasmid eluted with 100µl Nuclease-Free Water and collected by centrifuging at 13000rpm for 1 minute at room temperature.

### *Plasmid midipreparations using the Wizard Plus Midipreps DNA Purification System (Promega)*

Single colonies were picked from transformation plates and grown overnight in a 37°C shaking incubator in 100ml LB medium containing 50μg/ml ampicillin (Sigma). The culture was centrifuged at 6000rpm for 15 minutes in a Sanyo Harrier 18/80 benchtop centrifuge, the pellet drained well and completely resuspended in 3ml Cell Resuspension Solution by vortexing thoroughly. 3ml Cell Lysis Solution was added and mixed by inverting the tube until lysis was complete. Then 3ml Neutralisation Solution was added, the tube inverted 4 times, the sample split into 6 Eppendorf tubes and spun at 13000rpm for 15 minutes at 4°C. The cleared lysate was mixed with 10ml of resuspended Wizard Midipreps DNA Purification Resin, transferred to a Midicolumn and the resin/DNA mix bound to the column by applying a vacuum. To wash the DNA 2x15ml of Column Wash Solution was applied and drawn through by a vacuum. The resin was dried briefly by continuing the vacuum for 30 seconds after the solution had been pulled through. The column tip was spun at 13000rpm for 2 minutes at room temperature to remove any residual Wash Solution, 300μl preheated 65°C milliQ water was added, the column left to stand for 1 minute and the eluted DNA was collected by centrifuging at 13000rpm for 20 seconds at room temperature. To remove any residual resin fines the sample was spun at 13000rpm for 5 minutes, the supernatant transferred to a fresh tube and stored at -20°C.

### Gel extraction using QIAquick Gel Extraction Kit (Qiagen)

DNA fragments of 70bp-10kb were purified using the QIAquick Gel Extraction Kit (Qiagen).

The DNA band was excised from a 1% agarose gel with a razor blade, weighed and 3 volumes of Buffer QG per 1 volume of gel was added (100mg~100μl). The gel was dissolved by heating in a 50°C waterbath, 1 gel volume isopropanol added (unnecessary if DNA band was between 500bp and 4kb) and the sample transferred to a QIAquick spin column. After spinning the column at 13000rpm for 1 minute at room temperature, the flow through was discarded and the DNA washed with 750μl Buffer PE. After centrifugation at 13000rpm for 1 minute at room temperature, the flow through was discarded, the column respun for 1 minute and transferred to a fresh Eppendorf tube. To

elute the DNA 30µl of Nuclease-Free Water was added to the centre of the column, left to stand for 1 minute and the DNA finally collected by centrifugation at 13000rpm for 1 minute at room temperature.

## PCR purification using QIAquick PCR Purification Kit (Qiagen)

Single or double stranded DNA of 100bp-10kb from PCR and other enzymatic reactions were purified using the QIAquick PCR Purification Kit (Qiagen).
To 1 volume of PCR reaction 5 volumes of Buffer PB were added, the sample mixed, applied to a QIAquick spin column and spun at 13000rpm for 1 minute at room temperature. After the flow through was discarded, the DNA was washed with 750µl Buffer PE, the column centrifuged at 13000rpm for 1 minute at room temperature, the flow through discarded again and the column respun for 1 minute. Then the latter was transferred to a fresh Eppendorf tube, 30µl Nuclease-Free Water was applied to the centre of the column, left to stand for 1 minute and the eluted DNA collected by centrifuging at 13000rpm for 1 minute at room temperature.

## DNA sequencing

### *Sequencing by the Sanger Dideoxy-mediated chain termination method*

To sequence cloned HMG boxes the T7 Sequenase version 2.0 DNA sequencing kit (Amersham) was used according to the manufacturer's instructions.
Approximately 5µg of template DNA in 20µl milliQ water was denatured for 10 minutes at room temperature with 5µl freshly prepared 2M NaOH and purified with a minispin column (Pharmacia). 7µl of the denatured template was annealed to 1pM of either –40 M13 forward primer (supplied by the kit) or M13 reverse primer (Gibco BRL) and labelled with 5µCi $^{35}S$ following the manufacturer's instructions.
A 5% polyacrylamide gel (7.5M urea) was poured with the use of the Sequagel "Ultra Pure" Sequencing System (National Diagnostics), the entire 10µl sequencing reaction was loaded and run on a Sequi-Gen GT Nucleic Acid Electrophoresis Cell (BioRad) sequencing rig at 95-130W until the Xylene Cyanol dye front (light blue) reached about 10cm above the bottom of the plates. The gel was fixed for 15 minutes in a solution of 10%methanol, 10% glacial acetic acid, transferred to Whatman filter paper and dried on

a flatbed dryer using a BioRad Hydro Tech Vacuum Pump. An X-ray film was placed above the dried gel and exposed for several days at room temperature.

### Automated sequencing using BigDye terminator chemistry

The sequencing of plasmid constructs and PCR products to find mutant lesions in *sox10* were carried out by Dr. Paul Jones (Sequencing Core Facility, University of Bath) on an ABI DNA sequencer using BigDye terminator chemistry. For sequencing, typically 300-500ng plasmids were combined with 10pmoles of primer or alternatively, 3-10ng PCR product (approximately 500bp) with 5pmoles of primer in 6µl total volume. The resulting sequence files were viewed with the Chromas viewing programme.

## *Plasmids*

Cloning the full length *sox10* and parts thereof was accomplished using the pGEM-T vector (Promega) or pGEM-T Easy vector (Promega). Both T-vectors are based on the pGEM –5Zf(+) plasmid (Promega). pGEM-T was created by digesting the pGEM – 5Zf(+) plasmid with EcoRV and adding a 3'-terminal thymidine to both ends, whereas for pGEM-T Easy, the MCS (multiple cloning site) was modified by adding additional *Not* I and *Eco*RI restriction sites flanking the T-overhangs on each end.
Heatshock constructs were created in the pCSHSP vector (Halloran et al., 2000). It was created from the pCS2+ plasmid by replacing the sCMV promoter region with the heatshock promoter HSP70/4 (for a plasmid map see chapter 7, Figure 7.1).

## *Site-directed mutagenesis*

Site-directed mutagenesis was carried out following the instructions published in the online manual to the QuikChange Site-Directed Mutagenesis Kit (Stratagene, http://www.stratagene.com/manuals/200518.PDF).
Mutagenesis primers each complementary to opposite strands of the vector were designed according to the recommended guidelines to be 25-45 bases long with a Tm greater or equal to 78°C. The desired mutation site was located in the middle of the primers with 10-15 basepairs of correct flanking sequence to each side. The primers have a GC content of at least 40%, terminate with one or more GC clamps and were

PAGE purified. (For primer sequences see Primer design and sequence section in Materials and Methods).

Four 50µl reactions were set up, each containing either 5ng, 10ng, 20ng or 50ng of template (*cls*$^{m618}$ mutant heatshock construct to make the wild-type construct, wild-type to make the *cls*$^{tw2}$, *cls*$^{Y83X}$ and *cls*$^{E189X}$ mutant heatshock constructs), 1µl 125ng/µl of each primer (m618→WT1 and m618→WT2 or WT1→tw2 and WT2→tw2, WT1→Y83X and WT2→Y83X, WT1→E189X and WT2→E189X), 1µl 20mM (400µM final) dNTPs, 1µl (2.5u) Pfu Turbo DNA Polymerase (Stratagene), 5µl 10x reaction buffer and milli-Q water to make the reaction volume up to 50µl. A "no primer" control reaction was included containing 20ng of template, but no primers.

In an MJ Research PTC-DNA Engine reactions were subjected to 95°C for 30 seconds followed by 12 cycles 95°C for 30 seconds, 55°C for 1 minute and 68°C for 12 minutes 30 seconds (2 minutes/kb of plasmid length) if a point mutation was to be introduced (*cls*$^{m618}$→WT, WT→*cls*$^{tw2}$, WT→*cls*$^{Y83X}$). Alternatively, 16 cycles were employed when changing a single amino acid residue to create the *cls*$^{E189X}$ allele. 10µl of each reaction were analysed on a standard agarose gel for successful amplification. To the remaining 40µl reaction including the "no primer" control, 1µl (10u) *Dpn* I restriction enzyme (NEB) was added directly into the PCR tubes and samples incubated at 37°C for 60 minutes to digest the parental methylated strands. 1.5µl of those double-stranded nicked open circular plasmids containing the mutated site which gave the most and cleanest product and 1.5µl of the "no primer" control were transformed into Epicurian Gold Super-Competent cells (Stratagene) as described in the appropriate Material and Method section. Colonies were tested for successful mutagenesis by either restriction digest with *Tsp*RI or by sequencing across the mutated region. If the wild-type construct was created successfully, a *Tsp*RI restriction was gained compared to the *cls*$^{m618}$ mutant allele. To test this, a 20µl restriction digest with 2.5µl putative wild-type construct miniprep, 0.2µl BSA, 1µl *Tsp*RI (NEB), 2µl 10x NEB4 reaction buffer and 14.3µl milli-Q water was set up and incubated at 65°C for 3 hours. Positive clones were sequenced to ensure successful mutagenesis.

Since there was no similar test based on a change of restriction digest pattern available for the *cls*$^{tw2}$ heatshock construct, 350ng of several minipreps were directly sequenced across the mutagenesis site.

## *In situ protocols*

### Preparation of Dig and Flu labelled RNA *in situ* probes

Approximately 10μg of plasmid containing the gene of interest was linearised with 70-100u of restriction enzyme cutting 5' to the probe fragment in a large volume, typically 100μl. An aliquot of the reaction was loaded on a 1% agarose gel to check for complete digestion. The remaining sample was purified using the QIAquick PCR Purification Kit (Qiagen) and the concentration of the purified linearised plasmid estimated by comparison to a quantitative ladder (1kb ladder, NEB) on a 1% agarose gel. Probes were synthesised using the Dig RNA Labelling Kit (Boehringer/Roche). In a 20μl reaction 1-2μg of linearised plasmid was combined with 1μl 20x Dig-NTP or 2μl 10x Flu-NTP labelling mixture, 2μl 10x transcription buffer, 1μl RNAse inhibitor and 2μl RNA polymerase. Depending on the orientation of the cloned fragments T7 or SP6 RNA polymerase was used to create an antisense RNA probe. After incubation at 37°C for 2 hours, a 1μl aliquot was kept at –20°C. To the remaining reaction 1μl RNAse-free DNAse was added, the sample incubated at 37°C for 15 minutes, then 15μl 5M Ammonium acetate, 75μl absolute ethanol and 1.5μl seeDNA (Amersham) were added. The RNA was precipitated on ice briefly, spun at 13000rpm for 15 minutes at 4°C, the pellet washed with 500μl 70% ethanol, air dried for 1 minute after carefully removing all traces of ethanol and redissolved in 20μl fresh milliQ water. Again, a 1μl aliquot was removed and compared to the one kept earlier before DNAse treatment on a 1% agarose gel. Before storing the RNA probe at –80°C 80μl formamide (Sigma) were added. Typically, the *in situ* probes were used at a 1:200 working dilution made up in hybridisation mix.

### *In situ* hybridisation on whole mount zebrafish embryos

Digoxigenin or Fluorescein labelled RNA was detected with alkaline phosphatase conjugated anti-Dig or anti-Flu antibodies, respectively. To visualise labelled transcripts NBT/BCIP (Boehringer/Roche) or Fast Red (Boehringer/Roche) were used as colour substrates in the alkaline phosphatase reaction.

Embryos were staged according to Kimmel et al. (1995), dechorionated if they were older than 24hpf, anaesthetised and fixed in 4% paraformaldehyde/PBS [0.8% NaCl,

0.02% KCl, 0.02M $PO_4$ pH7.3] for at least 24 hours at 4°C. Embryos up to 24hpf were dechorionated after fixation. All further procedures were carried out at room temperature and using 1ml of solution at any time unless stated otherwise.
Embryos were dehydrated by washing 3x5 minutes in methanol, 1x10 minutes in methanol and then stored at –20°C for a minimum of 2 hours or until needed. Embryos were rehydrated by washing 5 minutes in 75% methanol/25% PBS, 5 minutes in 50% methanol/50% PBS, 5 minutes in 25% methanol/75% PBS and finally 4x5 minutes in PBT [0.1% Tween20 in PBS].
To improve the penetration of the probe, embryos between 18-somite and 22-somite stages were digested in a solution containing 10µg/ml proteinase K in PBT for 5 minutes, 24hpf embryos were treated for 15 minutes, 30hpf-48hpf embryos for 30 minutes and embryos up to 72hpf for 1 hour. Embryos were then refixed in 4% paraformaldehyde/ PBS for 20 minutes, washed 5x5 minutes in PBT and prehybridised at 65°C for 2-5 hours in hybridisation mix [50ml hyb mix contains 25ml formamide, 12.5ml 20xSSC, 0.5ml 5mg/ml heparin, 0.5ml 50mg/ml tRNA, 0.25ml 20% Tween20, 0.46ml 1M citric acid and 10.7ml milliQ water].
Embryos were hybridised in 200µl of a 1:200 dilution of the probe and incubated at 65°C overnight. The probe was saved and stored at –20°C for re-usage. To wash away non-specifically bound probe, embryos were first rinsed in hybridisation mix (HM), followed by 10 minute washes at 65°C with each of 75% HM/25% 2xSSC, 50% HM/50% 2xSSC, 25% HM/75% 2xSSC and 2xSSC in turn. Two high stringency washes at 65°C in 0.2xSSC for 30 minutes each were performed, followed by 10 minute washes at room temperature in each of 75% 0.2xSSC/ 25% PBT, 50% 0.2xSSC/50% PBT, 25% 0.2xSSC/75% PBT and PBT.
Embryos were blocked at room temperature in PBT/2% sheep serum/2mg/ml BSA for 2-5 hours and incubated with the appropriate antibody (anti-Dig or anti-Flu diluted 1:5000 in block solution) overnight at 4°C. The antiserum was discarded and the embryos rinsed once in PBT, then washed 6x15 minutes in PBT and 3x5 minutes in NBT/BCIP buffer [100mM TrisHCl pH 9.5, 50mM MgCl2, 100mM NaCl, 0.1% Tween20 in milliQ water] or Fast Red buffer [0.1M Tris pH8.0] depending on the colour substrate to be used in the detection step. NBT/BCIP stain solution was freshly prepared by dissolving one NBT/BCIP tablet (Boehringer/Roche) in 10ml milliQ water, whereas one Fast Red tablet (Boehringer/Roche) was crushed and dissolved in 2ml Fast

Red buffer. Embryos were stained in the dark up to 48 hours and the reaction then stopped by washing the embryos in PBT several times. To improve the optical clarity of stained embryos they were transferred to 50% glycerol/50% PBS a few hours prior to inspection under an MZ12 dissecting microscope (Leica). For a more detailed analysis and for documentation an Eclipse E800 microscope (Nikon) and DIC or fluorescence optics were used.

## Double *in situ* protocol

To detect expression of *sox10* and another gene of interest simultaneously a *sox10* antisense RNA probe was labelled with fluorescein and detected with an anti-Flu antibody coupled to alkaline phosphatase using Fast Red (Boehringer/Roche) as a colour substrate. Antisense RNA probes against *sox9a* and *sox9b* (kind gift of Bon-chu Chung, Institute of Molecular Biology, Taipei, Taiwan), *crestin* (kind gift of Paul Henion, Department of Neuroscience, Ohio State University, USA), *forkhead 6* (*fkd6*) and *dlx2* (kind gift of Craig Miller, Institute of Neuroscience, University of Oregon, USA) were labelled with digoxygenin, detected with an anti-Dig antibody also coupled to alkaline phosphatase, but using NBT/BCIP (Boehringer/Roche) as a purple colour substrate.

### *Preabsorption of anti-Flu antibody*

The anti-Flu antibody was found to give much less background when preabsorbed prior to usage. To preabsorb the antibody approximately 300 3-5dpf embryos were fixed in 4% paraformaldehyde in PBS at least overnight at 4°C, dehydrated in methanol as described in the single *in situ* protocol and kept at -20°C until needed. At the start of the double *in situ* protocol those embryos for preabsorption were treated alongside the experimental ones with the following modifications. Proteinase K treatment was carried out for 1 hour and instead of prehybridisation solution they were incubated in 1ml blocking solution containing 2µl anti-Flu antibody (1:500) at least overnight at 4°C.

### *Double in situ hybridisation on whole mount zebrafish embryos*

The protocol described previously for single *in situ* hybridisations was employed with the following modifications: The prehybridisation and hybridisation step was carried out at 66°C. The hybridisation mix contained both probes at a 1:200 dilution.
The blocking solution was modified to contain 5% sheep serum, 2mg/ml BSA and 1% DMSO in PBT. *sox10* was detected first with a 1:5000 dilution of preabsorbed anti-Flu antibody in blocking solution overnight at 4°C. The Fast Red colour reaction was stopped by 2 fast washes in PBT followed by 4x3 minute washes at room temperature in PBT. After developing the first colour reaction, bound anti-Flu antibody was removed by washing embryos 6x15 minutes in 100mM glycine, pH 2.2, 0.1% Tween20 at room temperature, rinsing 5x5 minutes in PBT, refixing in 4% paraformaldehyde/PBS for 30 minutes at room temperature and washing again 3x5 minutes in PBT before repeating the blocking step with the anti-Dig antibody. After the second colour reaction was stopped as described before for the first colour reaction, embryos were stored in 4% paraformaldehyde/PBS at 4°C.

## *Antibody staining protocols*

### Antibody staining with DAB as a substrate

Embryos up to 60hpf were stained using this protocol. Embryos to be antibody stained were dechorionated and fixed in 4% paraformaldehyde/PBS overnight at 4°C and then kept in PBSX at 4°C until needed. They were washed 2x5 minutes with PBSX [0.5% Triton-X (Sigma) in PBS], 4x30 minutes (2x30 minutes after *in situs*) in milliQ water, incubated in acetone for 7 minutes at –20°C and blocked for at least one hour in PBSDX [50ml 2xPBS pH7.3, 1g BSA (Sigma), 1ml DMSO (Sigma), 0.5ml 10% Triton X-100 and milliQ water up to 100ml] containing 15µl/ml horse serum (Sigma) block. Primary antibody, anti-Hu (Monoclonal antibody facility, University of Oregon) was diluted 1:5000 in PBSDX containing 15µl/ml horse serum and the embryos incubated overnight at 4°C. On the following day embryos were washed 4x30 minutes in PBSDX, the biotinylated secondary horse anti-mouse/rabbit antibody diluted 1:5000 in PBSDX with 15µl/ml horse serum and the embryos incubated overnight at 4°C.

The antibody was discarded, the embryos washed 4x30 minutes in PBSDX, incubated 45 minutes in the diluted ABC reagent (Vectastain ABC Elite kit), washed 3x30 minutes in PBSDX and 30 minutes in PBSX. The DAB substrate diluted in PBSX (14µl/ml) and the embryos incubated for 15 minutes. After transferring the embryos to a glass well plate 7µl hydrogen peroxide (Sigma) were added per 1ml diluted DAB solution. The stain was developed in the dark for 5 minutes up to a few hours. For long incubations it was necessary to add more hydrogen peroxide approximately every 10 to 15 minutes. The development was finally stopped by rinsing the embryos several times in PBSX and embryos were kept at 4°C.

## Antibody staining with a fluorescently labelled secondary antibody

This protocol was used for embryos older than 60hpf since the sensitivity of detection of structures deep within the tissue is greater with the fluorescent signal than with the DAB substrate.

Embryos were washed 3x5 minutes at room temperature in PBSX, 3x1 hour in milliQ water, preincubated in 0.75ml block [0.5% Triton X-100, 1% DMSO, 5% horse serum in PBSX] and incubated at room temperature overnight with the primary anti-Hu mouse antibody diluted 1:1100 in 0.75ml block.

The antibody was discarded, embryos rinsed in PBSX as described in the previous protocol, washed for 3x1 hour in PBSX and incubated at room temperature overnight with the secondary antibody, Alexa Fluor 546 rabbit anti mouse IgG (Molecular probes, Oregon), diluted 1:800 in 0.75ml block.

The antibody was discarded, embryos rinsed in PBSX, washed for 3x30 minutes in PBSX, transferred into 50% glycerol/PBS and incubated for 15 minutes. Embryos were analysed with an Eclipse E800 microscope (Nikon) using fluorescence microscopy (TRITC filter).

## Antibody staining as a secondary stain after a standard *in situ* hybridisation

If an antibody stain was to be carried out as a secondary stain after an *in situ* hybridisation, the following procedures of the *in situ* protocol described above were modified. Dechorionated embryos were fixed in 4% paraformaldehyde/PBS at 4°C for a

maximum of 24 hours. The duration of the proteinase K treatment on those embryos in the course of the *in situ* hybridisation was decreased to 15 minutes for 24hpf embryos, 25 minutes for up to 48hpf and 45 minutes for older stages). The *in situ* colour reaction product was slightly underdeveloped, so as not to mask any areas of weak antibody stain. At the end of the *in situ* protocol embryos were stored in PBT at 4°C ready for the antibody stain using either one of the two protocols described above.

## *Heatshock cDNA injection and rescue protocol*

Plasmid DNA to be injected was isolated in a midiprep using the Wizard Plus Midipreps DNA Purification System (Promega) and further cleaned by adding an equal volume of phenol:chloroform (1:1) and then, to the supernatant an equal volume of chloroform. A third of the volume 7.5M Ammonium acetate and 2.5 volumes absolute ethanol were added to precipitate the plasmid DNA followed by 1.5µl seeDNA (Amersham) to facilitate the pellet's visibility and tubes spun immediately at 13000 rpm at 4°C for 30 minutes. The pellet was washed with 1ml 70% ethanol, resuspended in 50µl milli-Q water and the concentration estimated by comparing several dilutions of the plasmid construct to a quantitative DNA ladder (NEB). Embryos from crosses of heterozygous *cls*$^{m618}$ were injected with 2-5nl 13ng/µl of the appropriate heatshock construct dissolved in milli-Q water at the 1-4 cell stage using a Nanojet II injector (Drummond Scientific Co.), incubated at 28.5°C at low density to recover and dead embryos removed after 6-8 hours. The injected volume was estimated by comparing the diameter of the bead of injected liquid with the diameter of a 1-cell stage embryo (approximately 700µm; Kimmel et al., 1995). Embryos were then heatshocked at 15-16 hpf by incubating embryos at 37°C for 1 hour. They were scored for rescue at 48 hpf using a MZ12 dissecting microscope (Leica). Melanophore rescue was defined as the presence of at least one melanophore of wild-type morphology.

## *Data analysis*

### Interpretation of sequencing data

The most likely identity of isolated HMG box sequences and RACE PCR clones was determined by a BLAST search on http://www.ncbi.nlm.nih.gov/BLAST/. Direct

comparison of the nucleotide or peptide sequence of two fragments was carried out with the use of the Pairwise BLAST analysis on http://www.ncbi.nlm.nih.gov/blast/bl2seq/bl2.html. Multiple sequence alignments were performed using the multiple sequence alignment program of the GCG package (Genetics Computer Group, Wisconsin package, Version 8.0-Unix, September 1994). Overlapping sequence fragments were put together in contigs using the following programs of the GCG package: gelstart, gelenter, gelmerge, gelassemble and geldisassemble. A consensus sequence was created using gelview.

## Statistical analysis of rescue experiments

The injection of heatshock cDNAs seemed to cause malformations in embryos of varying degree depending on the concentration of DNA injected and the time of heatshock. To determine the optimum amount of cDNA (13ng/µl) that could rescue the *cls* phenotype to the highest degree possible whilst giving the least malformations a range of different concentrations (6ng/µl, 13ng/µl, 25ng/µl and 60ng/µl each in 0.1% phenol red) of hs>$cls^{WT}$ cDNA were injected, each heatshocked at 2 different time points (early at 15-16hpf and late at 18-18.5hpf). The number of malformed wild-type and *cls* embryos was recorded at 48hpf and statistically analysed. First, a $\chi^2$ test (using the Yates' correction formula and one degree of freedom) was carried out on individual *cls*/WT pair datasets to determine whether the genotype influenced the degree of malformations. Since no significant difference was found at either 5%, 1% nor 0.1% level malformations did not vary with the genotype and hence for all subsequent analyses no distinction between *cls* and WT malformations were made.

A single factor anovar analysis with two levels was used to determine whether there was a difference in malformation when the hs>$cls^{WT}$ or the equivalent amount of control plasmid pCS (without *sox10* cDNA) was injected, whether malformation varied depending on the amount of cDNA injected, whether a heatshock increased malformations in general and whether the time of the heatshock was important. Finally, to test whether the presence of malformations had any effect on the probability of *cls* embryos being rescued, a $\chi^2$ test (using the Yates' correction formula and one degree of freedom) was employed.

# Chapter 3 – Cloning of a zebrafish *sox10* homologue

## *Introduction*

There are principally two approaches to identify the gene encoded by a mutant locus. Firstly by positional cloning which involves mapping the mutation to a critical interval on a chromosome followed by screening of Yeast artificial chromosome (YAC), Bacterial artificial chromosome (BAC) or P1-derived artificial chromosome (PAC) clones for open reading frames. The gene encoding the mutant locus can then be identified, for example by rescue experiments.
Alternatively, a candidate gene approach represents a more direct and potentially much faster method if a likely candidate gene is available. Here, this gene is cloned and tested by means of rescue experiments or presence of causative mutant lesions.
Since, as described in the Introduction chapter, a *sox10* homologue was considered a strong candidate for the *cls* gene, the candidate gene approach was chosen.

### Methods for cloning candidate genes

There are two approaches to clone such a candidate gene, by screening a cDNA library or by RT-PCR. Originally, both methods seemed equally suitable and hence were attempted in parallel.

At the time, there were only a few libraries available for zebrafish, but they included a λZAP cDNA library prepared from 19hpf zebrafish embryos (Appel and Eisen, 1998). This library has successfully been used to clone genes from zebrafish such as *nacre* (Lister et al., 1999). At 19hpf during zebrafish development, neural crest cells in the head and anterior trunk have entered the medial migration pathway, whereas crest cells in more caudal regions are still residing in a premigratory position. Both populations of cells in mouse embryos of equivalent stage would express *Sox10* (Herbarth et al., 1998; Southard-Smith et al., 1998). Furthermore, expression would also be expected at this stage in the developing otic vesicle in concordance with mouse and chick data (Southard-Smith et al., 1998; Cheng et al., 2000; Watanabe et al., 2000). Thus, we attempted to identify a *sox10* homologue from the 19hpf cDNA library using a

1.45kb fragment of mouse *Sox10* as a probe, kindly provided by Michelle Southard-Smith. This fragment has previously been used for a low stringency hybridisation demonstrating conservation of *Sox10* amongst multiple vertebrate species (Zooblot; Southard-Smith et al., 1999). The protocol is based on a hybridisation in Church buffer at 50-55°C overnight and 3x5 minutes washes in 2xSSC, 1%SDS at 55°C (Southard-Smith, M., personal communication).

Initially, this approach was attempted using the published Zooblot protocol. Since technical problems persisted even after thorough optimisation of conditions this method proved unsatisfactory in our hands and will not be considered further here.

The RT-PCR approach carried out in parallel by then appeared to be a more promising method and hence efforts were focused on this second approach. We took advantage of the high degree of conservation of *Sox* genes to attempt to clone at least a fragment of a zebrafish *sox10* by RT-PCR. Success with this approach required consideration of three major factors, the primer sequences, the PCR protocol and the stage of embryonic tissue as a source of mRNA.

A fragment of *Sox10* was first cloned from mouse using degenerate primers against conserved regions located just within the 5' and 3' boundaries of the HMG box (Wright et al., 1993). Michael Wegner's group then cloned a rat *Sox10* homologue using primers designed against almost identical sites first published by Yuan et al. (1995). 69% of HMG box fragments isolated from a primary rat Schwann cell culture showed high identity to *Sox10* (Kuhlbrodt et al., 1998a). Hence, we decided to attempt to isolate a zebrafish *sox10* homologue by RT-PCR using those degenerate primers and PCR protocol previously employed by Kuhlbrodt.

To prepare total RNA, we chose several stages during zebrafish development including 19hpf, 24hpf, 5dpf and adult head and skin. The rationale behind choosing those stages was based on the mouse S*ox10* expression pattern. In mouse, both premigratory and migrating neural crest cells express *Sox10* (Herbarth et al., 1998; Southard-Smith et al., 1998). In zebrafish embryos, cranial crest cells just caudal to the developing eye start to migrate subectodermally at 15-16hpf (Schilling and Kimmel, 1994). At 18hpf, some crest cells above somites 1-7 and anterior trunk crest cells have entered the medial migration pathway, whereas caudal to somite 9, the neural crest is in a premigratory stage. More posteriorly still, crest cells are still segregating from the neural keel (Raible et al., 1992). Thus, since neural crest cells of all premigratory and migratory phases are present, the 19hpf stage seemed an appropriate stage to use.

At 24hpf, neural crest cells are still entering the medial pathway along almost the entire length of the axis, whereas some crest cells in the region above somites 1-5 have started migrating on the lateral pathway (Raible et al., 1992). This stage appeared suitable since premigratory neural crest and crest cells migrating on both pathways were present. Cells on the lateral pathway give exclusively rise to pigment cells (Raible and Eisen, 1994), a cell-type severely affected in *cls* mutant embryos. Furthermore, counts of neural crest cells on the lateral pathway in *cls* mutant embryos show a 95% reduction in number compared to wild-type siblings indicative that *cls* is functioning by this stage (Kelsh et al., 1996).

Finally, since in mouse *Sox10* is maintained in certain crest derivatives and the developing ear during later stages, we chose two later stages as sources for RNA.

The 5dpf larval stage was chosen, because it represents the latest ethically acceptable larval stage to use, since larvae now start to actively feed and thus are protected under UK law.

In adult mice and humans, *Sox10/SOX10* is known to be expressed in Schwann cells in the peripheral nervous system (PNS) and in oligodendrocytes in the brain (Kuhlbrodt et al., 1998a; Kuhlbrodt et al., 1998b; Pusch et al., 1998). Hence, total RNA from adult brains was prepared. Finally, it was reported that melanocytes in mice express *Sox10* (Southard-Smith et al., 1998) and hence, adult zebrafish skin containing pigment cells was included.

The chosen strategy aimed to isolate the HMG box of zebrafish *sox10*, since this region is the most highly conserved. However, this general conservation across all *Sox* genes might make it very difficult to reliably identify a particular homologue across species. The sequence differences between two members of the same subgroup are very limited and hence, it would be unreliable to distinguish a *sox10* homologue from other subgroup E members based on the HMG sequence alone.

Thus, to confirm the identity of a suspected *sox10* fragment, we anticipated having to obtain a larger fragment outside the HMG box region employing the RACE-PCR (Rapid Amplification of cDNA Ends-Polymerase Chain Reaction) technique. Even with the full length clone in hand one would have to be careful in assigning an identity. Since fewer *Sox* gene sequences were available for comparison, a fair number of homologues published in Genbank were misclassified and only recently reassigned correctly. For example, the mouse gene originally published as a *Sox21* homologue (Tani et al., 1997) has now been confirmed to be a misnamed *Sox10* (Pusch et al., 1998). Likewise, trout

*SoxP1* reliably clusters with other known *Sox8* homologues. Thus, extensive phylogenetic analysis would be necessary to assign confidently an identity to any cloned *Sox* gene homologue.

## Results

### Cloning of a *sox10*-like HMG box from zebrafish by RT-PCR

Two RT-PCR strategies were considered to clone a *sox10*-like HMG box. Firstly, *sox10* specific primers could be designed. At the time, only mammalian *Sox10* homologues from human, mouse and rat were published in Genbank (Herbarth et al., 1998; Kuhlbrodt et al., 1998a; Pingault et al., 1998a; Southard-Smith et al., 1998). Conserved regions identified between the above mammalian genes might not necessarily show the same degree of conservation in fish. Hence, choosing suitable sites for primers with acceptable degeneracy was expected to be very unreliable and difficult.
The second strategy would involve amplification with primers in regions conserved across all *Sox* genes and subsequently identify a *sox10* homologue from the mixture. The use of such generic *Sox* gene primers had proven successful before in isolating *Sox10* homologues from mouse and rat. Thus, this strategy seemed more promising and hence was chosen to clone a *sox10*-like HMG box from zebrafish by RT-PCR.
These primers, 5' Sox10 and 3'Sox10, (for sequences see Materials and Methods; Yuan et al., 1995) amplify a mixture of HMG box sequences from cDNA corresponding to *Sox* genes expressed at any particular stage during development (Figure 3.1). Fragments were then cloned, sequenced and aligned with previously characterised members of this gene family to identify a likely *sox10* homologue.

#### *Isolation of HMG box sequences cloned by RT-PCR from cDNA of 4 different developmental stages*

Total RNA was prepared from 19hpf, 24hpf, 5dpf and adult brain and skin of AB wild-type zebrafish. The rationale behind choosing these stages has been described in detail in the Introduction to this chapter. RNA samples of each stage were quantified spectrophotometrically by measuring the absorption of a diluted aliquot of RNA at 260nm wavelength (Table 3.1). A second spectrophotometric reading at 280nm

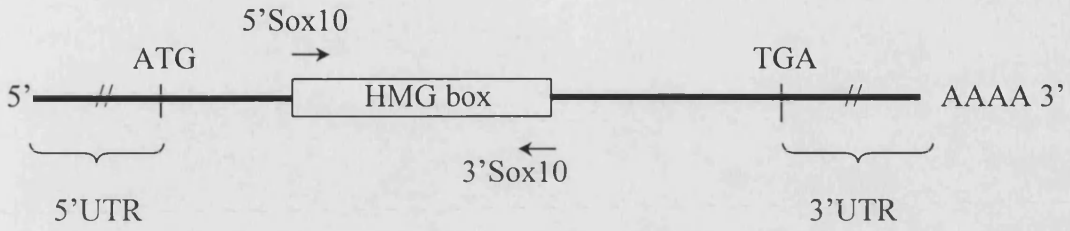



**Figure 3.1:** The general structure of a *Sox10* homologue.

Start (ATG) and Stop (TGA) codons mark the beginning and the end of the coding region. The latter is preceeded by a 5' untranslated region (5'UTR) and followed by a 3' untranslated region (3'UTR). The degenerate primers, 5'Sox10 and 3'Sox10, are located just within the HMG box and were used to amplify HMG box sequences from various stages by RT-PCR.

| Stage | 19hpf | 24hpf | 5dpf | Adult tissues |
|---|---|---|---|---|
| $A_{260nm}$ | 0.062 | 0.156 | 0.258 | 0.019 |
| $A_{280nm}$ | 0.042 | 0.076 | 0.111 | 0.011 |
| $A_{260nm}/A_{280nm}$ | 1.48 | 2.05 | 2.32 | 1.73 |
| RNA conc. | 740ng/µl | 1.87µg/µl | 3.10µg/µl | 760ng/µl |
| µg RNA used per RT reaction | 1µg | 1µg | 2µg | 1µg |
| µl cDNA used in PCR | 2µl | 2µl | 2µl | 2µl |
| ng PCR product cloned | ~200ng | ~100ng | ~100ng | ~350ng |
| Total white colonies | 64 | ~150 | ~100 | 83 |
| Colonies grown up | 19 (29.7%) | 18 (~10%) | 20 (~20%) | 29 (34.9%) |
| Positive clones | 18 (94.7%) | 15 (83.3%) | 13 (65.0%) | 29 (100%) |
| Clones sequenced | 10 (55.6%) | 12 (80%) | 12 (92.3%) | 13 (44.8%) |
| *Sox* genes | 9 (90%) | 10 (83.3%) | 6 (50%) | 11 (84.6%) |
| No homol. to *Sox* genes | 1 (10%) | 2 (16.7%) | 6 (50%) | 2 (15.4%) |

**Table 3.1:** Summary of results from cloning of HMG boxes by RT-PCR using generic *Sox* gene primers

For each developmental stage, this table lists spectrophotometric measurements of absorption of 260nm and 280nm wavelength light by total RNA samples. The concentration of each sample was calculated using the $A_{260nm}$ reading. The ratio $A_{260nm}/A_{280nm}$, ideally close to 2.0, provides an estimate for purity of the RNA sample. The amounts of µg total RNA included in each reverse transcription (RT) reaction, the numbers of µl cDNA reaction out of a total of 200µl used per PCR reaction and the amount of PCR product in ng, estimated by agarose gel electrophoresis, which was ligated into 25ng pGEM-T vector, are listed. In each case, the entire ligation reaction was transformed. The table summarises the number of white colonies obtained from each cloning experiment and how many of those were chosen for plasmid minipreparations. It also lists the number of constructs which contained inserts of the expected size (positive clones), how many of those were sequenced and the proportion of HMG box sequences identified from each stage. conc., concentration; homol., homology.

wavelength gave an estimation of protein contamination. When comparing the two readings $A_{260}/A_{280}$ the ratio was usually close to 2.0, typical for RNA of high purity. (Table 3.1). Gel electrophoresis further showed that the RNA was of good quality. This was judged by two criteria, presence of three strong bands corresponding to the 3 ribosomal rRNA types (23S, 16S and 5S) and a slight smear indicative of mRNA transcripts (Figure 3.2A, 3.4A and B). Total RNA isolated from adult tissues contained a lot more genomic DNA compared to the other stages (Figure 3.4A). Nevertheless, this RNA was used, because the short extension time during the amplification reaction would likely bias against a longer genomic product, if an intron disrupted the HMG box. Even if such a product was obtained it could be distinguished by its larger size due to the intronic insertion. To control for amplification from genomic DNA, a reverse transcription reaction without the reverse transcriptase was set up alongside ("-RT"). Indeed, no product was obtained in the subsequent RT-PCR reaction even with 2µl of the "-RT" control reaction (Figure 3.4C).

From each stage, 1-2µg of total RNA (Table 3.1) were reverse transcribed as described in the Materials and Methods. Six equivalent RT-PCR reactions were set up with a suitable amount of cDNA, usually 2µl, previously determined by a pilot RT-PCR. Increasing the amount of cDNA also increased the amount of RT-PCR product (Figure 3.3A and 3.4C). The rationale behind setting up 6 reactions was to allow pooling of the PCR product so as to have sufficient purified product for cloning. All experimental reactions amplified the expected single product of 204bp, which was never observed in the "no cDNA" control (-ve lane) for 24hpf and 5dpf stages (Figure 3.2B, 3.3A). A barely visible band was observed in the "no cDNA" control for the 19hpf stage and adult tissues (Figure 3.4C). Most likely this indicates a DNA contamination whilst setting up this control reaction or a very slight DNA contamination of one or both primers. However, bands derived from experimental reactions were much stronger in comparison. Control reactions containing only one of the primers ("5'Sox10 only" and "3'Sox10 only") or none of the primers ("no primers") were included in RT-PCRs from adult tissues and 19hpf stage. As expected, no product was amplified in these control reactions (Figure 3.4C).

The remainder of all PCR reactions were pooled, purified using the QIAquick PCR Purification Kit (Qiagen) and directly ligated into the pGEM-T vector. Constructs were transformed into $CaCl_2$ competent cells by a standard heatshock protocol and cells

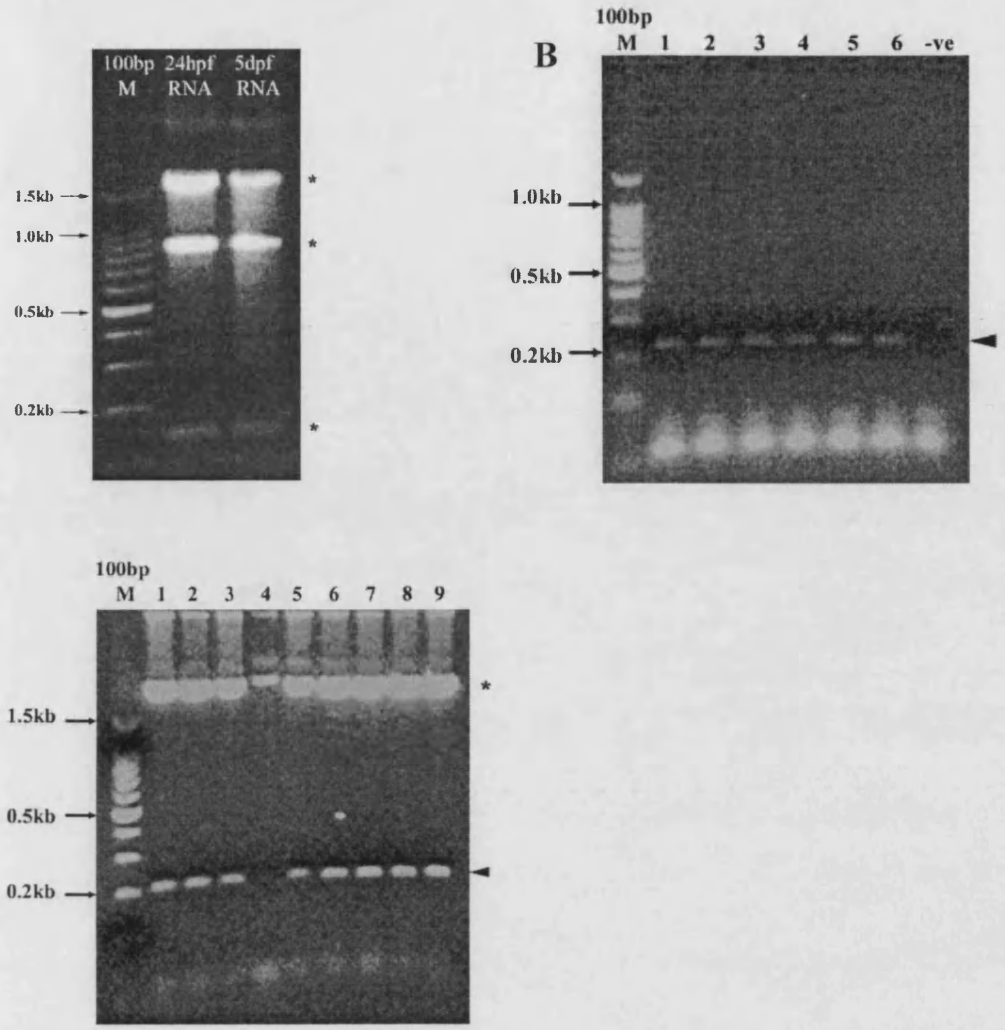



**Figure 3.2:** RT-PCR from 5dpf zebrafish embryos using generic Sox primers
(A) The quality of total RNA from 5dpf larvae and 24hpf zebrafish embryos was good as judged by the presence of the 3 ribosomal RNA bands, 23S, 16S and 5S (asterisk), in combination with a slight background smear corresponding to mRNA transcripts. 100bp M, 100bp marker (Promega). (B) Samples of the RT-PCR reactions are shown. A single specific product (arrowhead) of the expected size (204bp) was obtained in all experimental reactions (lanes 1-6), but not in the "no cDNA" control (-ve lane). These 204bp products corresponded to HMG box sequences. Sizes were judged by comparison to a 100bp marker (100bp M, Promega). Concentration and purity was estimated by spectrophotometric determination of absorption at 260nm and 280nm. (C) Samples of plasmid DNA minipreparations were digested with *Bam*HI to test the size of the insert. Here, 9 such digests are shown (lane 1-9), together with a 100bp marker (100bp M, Promega). The vector band (asterisk) and inserts of the correct size (arrowhead) are marked.

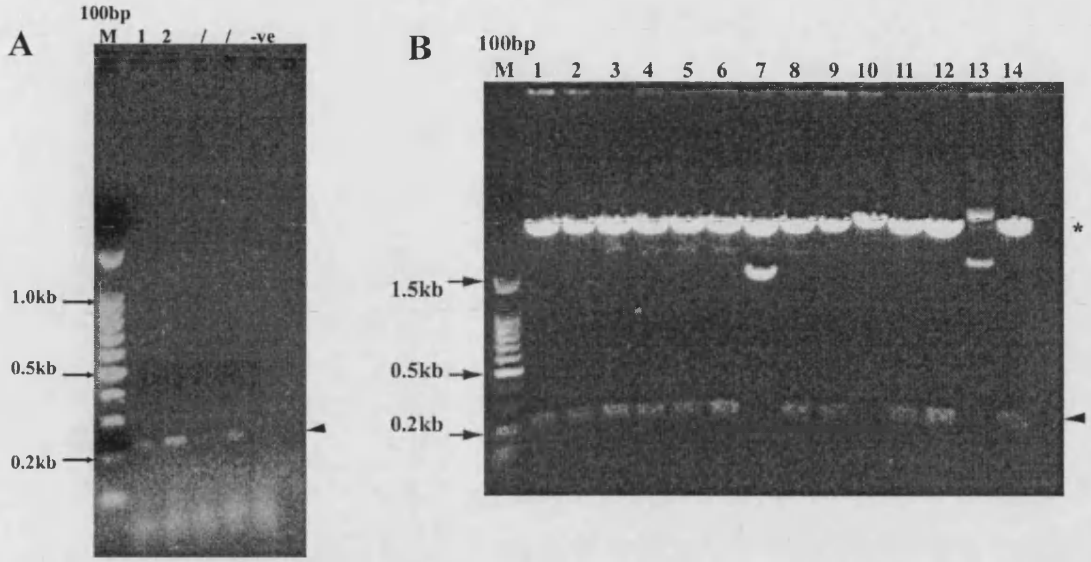
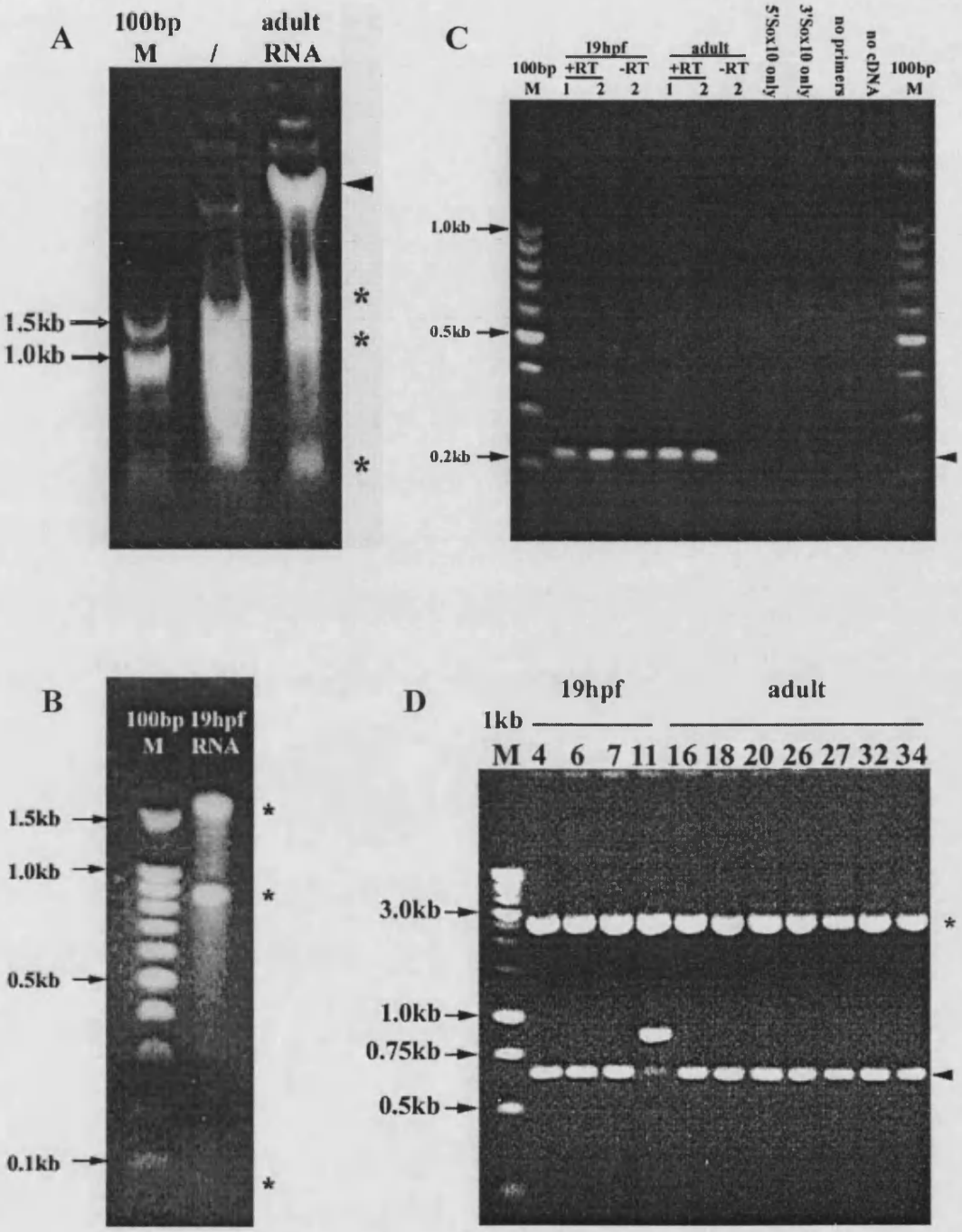



**Figure 3.3:** RT-PCR from 24hpf zebrafish embryos using generic Sox primers (A) The amounts of cDNA were tested in RT-PCR reactions with 2µl cDNA (lane 2) producing a stronger band of the correct size (arrowhead) than with 1µl cDNA (lane 1). No such band was observed in the "no cDNA" control (-ve lane). To judge band sizes, a 100bp marker (100bp M, Promega) was run alongside experimental samples. Lanes marked with a dash, /, are irrelevant to this experiment. (B) Plasmid minipreparations were digested with *Bam*HI to test for presence and size of inserts. All, but clone 7 and 13 contain inserts of the expected 204bp size (arrowhead). The size of the vector (asterisk) and insert bands were judged by comparison to a 100bp marker (100bp M, Promega).

**Figure 3.4:** RT-PCR from 19hpf zebrafish embryo and adult tissues using generic Sox primers

(A, B) The quality of total RNA from heads and skins of 9 adults (A) and 19hpf stage embryos (B) was good as judged by the presence of the 3 ribosomal RNA bands, 23S, 16S and 5S (asterisk), in combination with a slight background smear corresponding to mRNA transcripts. A large amount of genomic DNA (arrowhead) was present in the RNA preparation from adult tissues (A). The lane marked with a dash, /, is irrelevant to this experiment (A). A rough size estimate was provided by a 100bp marker (100bp M, Promega). (C) Samples of RT-PCR reactions and controls are shown alongside a 100bp marker (100bp M, Promega). RT reactions were set up with RNA from 19hpf embryos and adult tissues. RT-PCR reactions (+RT lanes) using 2µl cDNA (lanes "2") instead of 1µl cDNA (lanes "1") produced stronger bands of the expected 204bp size (arrowhead). Control reactions for the reverse transcription step are labelled "-RT", because no RT enzyme had been included. As expected, no band was observed in the "-RT" control reaction from adult tissues (adult, -RT). However, a strong product was obtained in the equivalent control reaction from the 19hpf stage (19hpf, -RT), which is most likely derived from an amplification off genomic DNA. Control reactions for the RT-PCR amplification included reactions with only one of the primers (lanes "5'Sox10 only" and "3'Sox10 only"), none of the primers (lane "no primers") and no cDNA (lane "no cDNA"). As expected, no products were observed in single primer only and no primer controls. However, a faint band was seen in the "no cDNA" control, which might be due to slight contamination of a primer solution or any other PCR reagent. (D) Samples of plasmid DNA minipreparations were digested with *Pvu*II to test the size of the inserts. Here, 4 such digests from the 19hpf stage and 7 from the adult tissue cloning are shown. All but clone 11 contain an insert of the expected 650bp size (arrowhead) compared to a 1kb marker (1kb M, Promega). Clone 4, renamed *SoxS4*, was later shown to exhibit high percentage nucleotide identity to *sox10* homologues.


A
100bp
M
adult
RNA
1.5kb
1.0kb
C
19hpf
adult
100bp
+RT
-RT
+RT
-RT
5'Sox10 only
3'Sox10 only
no primers
no cDNA
100bp
M
1 2 2 1 2 2
1.0kb
0.5kb
0.2kb
B
100bp 19hpf
M RNA
1.5kb
1.0kb
0.5kb
0.1kb
D
19hpf
adult
1kb
M 4 6 7 11 16 18 20 26 27 32 34
3.0kb
1.0kb
0.75kb
0.5kb

plated on LB plates containing ampicillin, isopropylthiogalactoside (IPTG) and 5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside (X-gal) for blue/white selection. The pGEM-T vector contains a gene conferring ampicillin resistance to cells harbouring the plasmid and thus enabling them to grow on ampicillin plates. Usually, 60-150 white colonies likely to contain plasmids with inserts were obtained per transformation and a convenient number (10-20 colonies) picked for plasmid minipreparations using the alkaline lysis method (Materials and Methods, Table 3.1). Preference was given to those colonies growing amongst blue ones indicating that the blue/white selection was working in that area of the plate. Furthermore, very small white colonies grown at high densities were avoided since they might represent satellite colonies. These false positives do not contain the plasmid and can only grow in ampicillin depleted zones found in close proximity to ampicillin resistant clones.

Plasmid constructs derived from 24hpf and 5dpf stage RNA were analysed by restriction digestion with *Bam*HI (Promega). There are *Bam*HI recognition sequences engineered into each primer, 5'Sox10 and 3'Sox10. Digestion of a positive clone with this enzyme releases the ligated HMG box fragment, which is recognised as a band of 204bp on an agarose gel together with a 3018bp band, corresponding to the remaining vector (Figure 3.2C and 3.3B). Of 18 white colonies grown up from the 24hpf stage, 15 (83.3%) contained an insert of the correct size. This was also true for 13 out of 20 (65%) white colonies derived from the 5dpf stage.

Analysis of clones from the 19hpf and adult stage was performed in a similar way, except that, for trivial reasons, an alternative restriction enzyme, *Pvu*II (Promega), was used to excise the insert. *Pvu*II recognition sites are located more distantly from the integration site and thus release an insert 436bp bigger than the cloned fragment. Thus, positive clones are recognised by a 640bp band in combination with a 2560bp band corresponding to the remaining vector (Figure 3.4D). All of 29 (100%) constructs derived from adult tissues and 18 out of 19 (94.7%) from the 19hpf stage contained an insert of the correct size.

12 positive clones from each of the 24hpf and 5dpf stages, plus 10 from 19hpf and 13 from adult stages were manually sequenced. Preference was given to clones digested to completion, which might indicate a cleaner plasmid minipreparation. Furthermore, constructs releasing additional bands were avoided. Preliminary assignment of sequence homologies of positive clones to published *Sox* genes were determined by a BLASTN search of the Genbank database (Table 3.2). Vector and primer sequences were removed

prior to the alignment. The quality of sequence of one insert cloned from the 5dpf stage was insufficient for analysis and 5 additional clones did not show any homology to *Sox* genes outside the primer sites. Two clones from adult tissues and the 24hpf stage and one clone from the 19hpf stage were also discarded for the same reason. One clone including its primer sites isolated from the 19hpf stage showed a high percentage nucleotide identity (119 out of 146bp, 81.5%) to rat *Sox10.* This clone, *SoxS4,* was thus resequenced 3 times to improve sequence quality. The consensus sequence was then used for further investigations.

### *Sequence analysis of the putative sox10-like HMG box clone SoxS4*

Although very similar to *Sox10*, we wanted to examine whether *SoxS4* might be a *sox9* or *sox8* homologue. This concern is justified, since all 3 genes show high similarities throughout the HMG box. To address this question we made sequence comparisons based on the nucleotide sequence using BLASTN and based on the predicted protein sequence using BLASTX (Figure 3.5). Both results were suggestive of a *sox10* identity, since *SoxS4* matched *Sox10* homologues slightly better than *Sox9* homologues, *Sox10*'s closest relative. In sequence alignments with HMG domains of known members of the subgroup E, SoxS4 was 94.1% identical to Sox10, but only 92.6% identical to Sox9 and 91.2% identical to Sox8 homologues (Figure 3.6). The HMG domain has been defined by convention as the stretch of 80 amino acids starting with the highly conserved residues PHVKRP and ending with residues RRRKNG in Sox10 homologues and RRRKSV in Sox8 and Sox9 homologues (Kuhlbrodt et al., 1998a; Pusch et al., 1998; Bowles et al., 2000). The amino acid identities in the alignments with clone SoxS4 and other known Sox proteins were calculated in the overlapping region. Primer sequences (MNAFMVW and PDYKYQP) used to identify SoxS4 were included in the alignment since the amino acid sequence they encode is fully conserved across Sox10 proteins and very highly conserved in all known Sox proteins (Bowles et al., 2000). This assumption was later confirmed to be true for *SoxS4* when the full cDNA had been cloned and sequenced (Figure 3.17). Hence, amino acid identities were calculated between residue 7 (M, methionine) and 74 (P, proline) (Figure 3.6).

To further test the identity of clone SoxS4, its sequence was aligned with the HMG domains of the two zebrafish paralogues *sox9a* and *sox9b*, which were available to us prior to publication (Chiang et al., 2001; Figure 3.6). This analysis clearly showed that SoxS4 was equally similar to both zebrafish Sox9 sequences, and less related to

**Table 3.2:** Composition of *Sox* HMG box sequences derived from different stages by RT-PCR reactions using generic *Sox* primers.

The table summarises tentative identities of isolated *Sox*-like sequences. RT-PCR reactions were set up from total RNA derived from 19hpf and 24hpf whole zebrafish embryos, whole 5dpf larvae and adult tissues (brain and skin). The first two columns describe the number of sequenced clones showing homologies to *Sox* genes and their proportion of total *Sox*-like clones. Column 3 and 4 describe the best matching sequence to the isolated HMG box fragment obtained from a BLAST search and the extent of sequence identity. Numbers in brackets represent the length of the isolated fragment after vector and primer sequences were removed. The ratio corresponds to the number of nucleotides of the isolated clone which are identical to the best matching sequence. For example, one 178bp clone isolated from adult tissue exhibited 155 nucleotides identical to zebrafish *sox11b* over a region of 161bp of the *sox11b* sequence. The E value in column 5 decreases exponentially with the significance of a match. Hence, with an E value of 1 assigned to a BLAST result, one might expect to see one sequence match with a similar score just by chance. It thus represents a probability that the tentative identity assigned in column 3 is a true match. Finally, the last column assigns the most likely identity of the isolated clone to a zebrafish homologue.

*LG27*, an Eublepharis macularius (leopard gecko) derived *Sry*-related sequence, and the *SoxB2* from Strongylocentrotus purpuratus (sea urchin) have not been renamed according to the *Sox* gene nomenclature. However, *LG27* shows similarities to mammalian *Sox4* homologues and *SoxB2* to Xenopus *Sox2*.

One clone derived from the 19hpf stage showed high similarity to *Sox10* homologues. It was renamed *SoxS4* and investigated further.

*Chelydra serpentina* (*Chelydra*), Snapping turtle; *Danio rerio* (*D. rerio*), zebrafish; *Ovis aries* (*O. aries*), sheep; *Strongylocentrotus purpuratus* (*S. purpuratus*), sea urchin; *Xenopus laevis* (*Xenopus*), African clawed frog. nucl., nucleotide; id., identity

| stage | number of clones | % *Sox*-like clones | best matching sequence | Nucl. id. to most similar sequence (total length) | Expect value (E) | likely identity |
|---|---|---|---|---|---|---|
| **Adult tissue** | 5 | 55.6% | mouse *Sox19* | 97/116 (174)<br>84/101 (142)<br>65/77 (149)<br>61/74 (191)<br>41/47 (169) | $5e^{-13}$<br>$7e^{-09}$<br>$4e^{-04}$<br>0.53<br>0.007 | *Sox19* |
| | 1 | 11.1% | *D. rerio sox11b* | 155/161 (178) | $3e^{-64}$ | *Sox11b* |
| | 1 | 11.1% | mouse *Sox13* | 75/84 (190) | $1e^{-17}$ | *Sox13* |
| | 1 | 11.1% | human *SOX4* | 25/27 (135) | 1.4 | *Sox4* |
| | 1 | 11.1% | *LG27* | 45/54 (141) | 0.37 | *Sox4?* |
| **5dpf** | 3 | 50% | *D. rerio sox11b* | 113/116 (137)<br>72/74 (108)<br>65/73 (142) | $2e^{-46}$<br>$6e^{-27}$<br>$2e^{-05}$ | *Sox11* |
| | 1 | 16.7% | human *SOX4* | 99/115 (115) | $2e^{-08}$ | *Sox4* |
| | 1 | 16.7% | human *SOX6* | 43/44 (202) | $6e^{-13}$ | *Sox6* |
| | 1 | 16.7% | *D. rerio sox31* | 18/18 (188) | 9.1 | *Sox31* |
| **24hpf** | 6 | 60% | *D. rerio sox19*<br><br>mouse *Sox19* | 131/139 (142)<br>36/36 (109)<br>125/147 (197)<br>121/146 (196)<br>70/83 (136)<br>35/38 (131) | $2e^{-49}$<br>$8e^{-11}$<br>$7e^{-22}$<br>$2e^{-16}$<br>0.001<br>$9e^{-05}$ | *Sox19* |
| | 2 | 20% | *D. rerio sox11a*<br>*Xenopus Sox11* | 175/177 (201)<br>57/62 (135) | $7e^{-90}$<br>$6e^{-12}$ | *Sox11* |
| | 1 | 10% | *O. aries Sox2** | 123/148 (219) | $4e^{-11}$ | *Sox2* |
| | 1 | 10% | *D. rerio sox31* | 132/136 (181) | $6e^{-56}$ | *Sox31* |
| **19hpf** | 2 | 22.2% | *D. rerio sox21* | 128/131 (189)<br>108/114 (118) | $3e^{-55}$<br>$5e^{-34}$ | *Sox21* |
| | 2 | 22.2% | mouse *Sox13* | 112/134 (192)<br>92/110 (197) | $2e^{-16}$<br>$2e^{-09}$ | *Sox13* |
| | 1 | 11.1% | rat *Sox10* | 119/146 (198) | **$1e^{-27}$** | ***Sox10*** |
| | 1 | 11.1% | *D. rerio sox11b* | 146/152 (151) | $4e^{-63}$ | *Sox11b* |
| | 1 | 11.1% | *Chelydra Sox5* | 32/35 (125) | 1.3 | *Sox5* |
| | 1 | 11.1% | Human *SOX6* | 32/33 (148) | $2e^{-06}$ | *Sox6* |
| | 1 | 11.1% | *S. purpuratus SoxB2** | 22/23 (126) | 1.3 | *SoxB2* |

```
Sequences producing significant alignments:                          Score

sp|O09141|SX21_MOUSE  TRANSCRIPTION FACTOR Sox-21 (TRANSCRIPTION...   120
emb|CAA04485|  (AJ001029) Sox10 protein [Rattus norvegicus]           120
gi|3264586  (AF047043) Sox-10 protein [Mus musculus]                  120
emb|CAA04576|  (AJ001183) SOX10 protein [Homo sapiens]                120
gi|3264588  (AF047389) Dominant megacolon mutant Sox-10 protein ...   120
gi|2826523  (AF017182) putative transcription factor; Sox10 [Mus...   120
sp|P48434|SOX9_CHICK  TRANSCRIPTION FACTOR Sox-9 gi|1589736 (U1...    118
dbj|BAA25296|  (AB012236) Sox9 [Gallus gallus]                        118
sp|O18896|SOX9_PIG  TRANSCRIPTION FACTOR Sox-9 gi|2554931 (AF02...    118
sp|P48436|SOX9_HUMAN  TRANSCRIPTION FACTOR SOX-9 gi|1082721|pir...    118
gb|AAD17974|  (AF106572) Sox9 [Alligator mississippiensis]            118
pir||S52469  Sox9 protein - mouse                                     118
dbj|BAA24365|  (AB006448) Sox9 [Oncorhynchus mykiss]                  116
gi|3126870  (AF061784) Sox9 homolog [Calotes versicolor]              116
pir||S68425  SoxP1 protein - rainbow trout gi|1199770|dbj|BAA11...    115
sp|Q04888|SX10_MOUSE  Sox-10 PROTEIN gi|423573|pir||S30242 sox1...     99
gi|1575717  (U70441) Sox10 [Mus musculus]                              99
sp|Q04887|SOX9_MOUSE  TRANSCRIPTION FACTOR Sox-9 gi|423577|pir|...     96
sp|Q04886|SOX8_MOUSE  Sox-8 PROTEIN gi|423576|pir||S30246 sox8 ...     94

sp|O09141|SX21_MOUSE TRANSCRIPTION FACTOR Sox-21, TRANSCRIPTION FACTOR Sox-M
                     gi|1872475 (U66141) transcription factor Sox-M [Mus
                     musculus] Length = 533
Score =  120 bits (298), Expect = 1e-27
Identities = 59/66 (89%), Positives = 62/66 (93%), Gaps = 2/66 (3%)
Frame = +1

SoxS4: 1   MNAFM-FGQAGR-KLRDQYPHLHNAELSKTLGKLWRLLNETDKRPFIEEAERLRKQHKKD 174
           MNAFM + QA R KL DQYPHLHNAELSKTLGKLWRLLNE+DKRPFIEEAERLR QHKKD
Sox21: 202 MNAFMVWAQAARRKLADQYPHLHNAELSKTLGKLWRLLNESDKRPFIEEAERLRMQHKKD 261

SoxS4: 175 YPDYKYQP 198
           +PDYKYQP
Sox21:  262 HPDYKYQP 269
```

**Figure 3.5:** The preliminary SoxS4 sequence aligns with subgroup E Sox proteins in a BLASTX search.

BLASTX compares the given nucleotide sequence in all six reading frames to peptide sequences in the database. Even the preliminary SoxS4 sequence obtained from the first read shows higher amino acid identities to known Sox10 than Sox9 or Sox8 homologues. The mouse Sox21 sequence is homologous to and now recognised as a mislabelled Sox10 (Pusch et al., 1998). Subsequently, the SoxS4 fragment was resequenced several times on both strands, which resolved all mismatches to mouse Sox21 in the 16 N-terminal residues of the fragment.

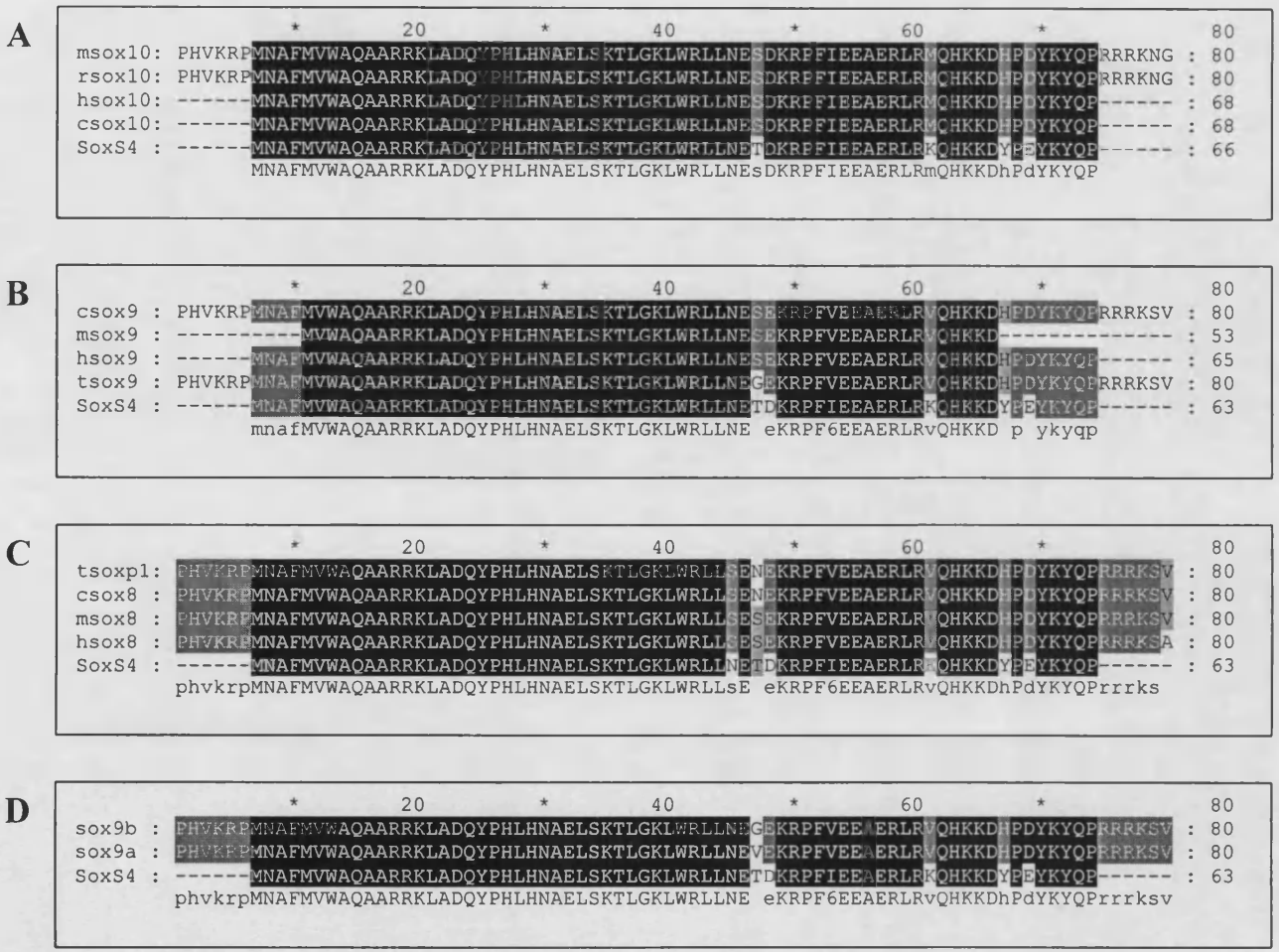



**Figure 3.6:** Alignments of the confirmed *SoxS4* sequence with published members of subgroup E

Multiple sequence alignments were produced with Pileup (gcg Package, see Materials and Methods) and regions of identity shaded in black (Genedoc). Amino acid identities were calculated between residues 7-74.(A) Pileup of *Sox10* homologues reveals slightly higher identity of *SoxS4* to this group (94.1%) than other closely related *Sox* genes. (B) Pileup of *SoxS4* with *Sox9* homologues (92.6% identity). (C) Pileup of *SoxS4* with *Sox8* homologues (91.2% identity). (D) Pileup of *SoxS4* with zebrafish *sox9a* and *sox9b* (92.6% identity) showing that *SoxS4* is indeed not a zebrafish *sox9*, the closest relative to *sox10*. It is noteworthy that the tentative identity of clone *SoxS4* as a *sox10* homologue rests on the conserved residue D at position 48. c, chicken; h, human; m, mouse; r, rat; t, trout.

either of them than they were to each other. Thus, this result argued against the possibility that *SoxS4* was a *sox9* homologue and further supported the identity of *SoxS4* as a *sox10* clone.

## Cloning of a full length *sox10* homologue by RACE PCR

SoxS4 showed great promise as a true *sox10* homologue, but it was necessary to extend the analysis to a full length cDNA. It was important to demonstrate a high degree of conservation outside the DNA binding domain. Furthermore, this larger fragment could also be used as an *in situ* hybridisation probe. If the cloned gene represented a *sox10* homologue, its expression pattern would be expected to be reminiscent of mouse *Sox10* (Chapter 4).

### *RACE-PCR cloning strategy*

The common approaches to cloning a full length cDNA are either screening a cDNA library or extending a previously cloned fragment of the gene of interest by RACE PCR (Rapid Amplification of cDNA Ends-Polymerase Chain Reaction). The RACE PCR approach was chosen, since difficulties were experienced previously with library screens.

We chose the widely used SMART RACE PCR Kit (Clontech), which uses the following strategy to clone a 5'RACE product (Figure 3.7). First strand synthesis is primed using an oligo(dT) based primer (5'cDNA Synthesis Primer, 5'CDS). After extending this primer, the Superscript II reverse transcriptase adds a short sequence of cytosines (dC) to the 3' end of the first cDNA strand. The SMART oligo supplied in the kit anneals to this region and thus serves as an extended template for the reverse transcriptase during its subsequent synthesis of the antisense cDNA strand. This is important, because DNA polymerases in PCR reactions require a primer presenting a 3' hydroxyl group for DNA synthesis. Without the SMART oligo extension, the far 5'end of the gene would be lost in subsequent cycles of PCR amplification.
During the first round of 5' RACE PCR, the RNA/DNA hybrid is denatured and the 3' end of the Long Universal Primer (Long UP) anneals to the SMART oligo sequence incorporated into the first cDNA strand. This results in a double stranded cDNA template. Subsequent gene specific rounds of 5'RACE PCR amplify the region between a Short Universal Primer (Short UP), which binds to the 5' end of the Long UP and a 5'

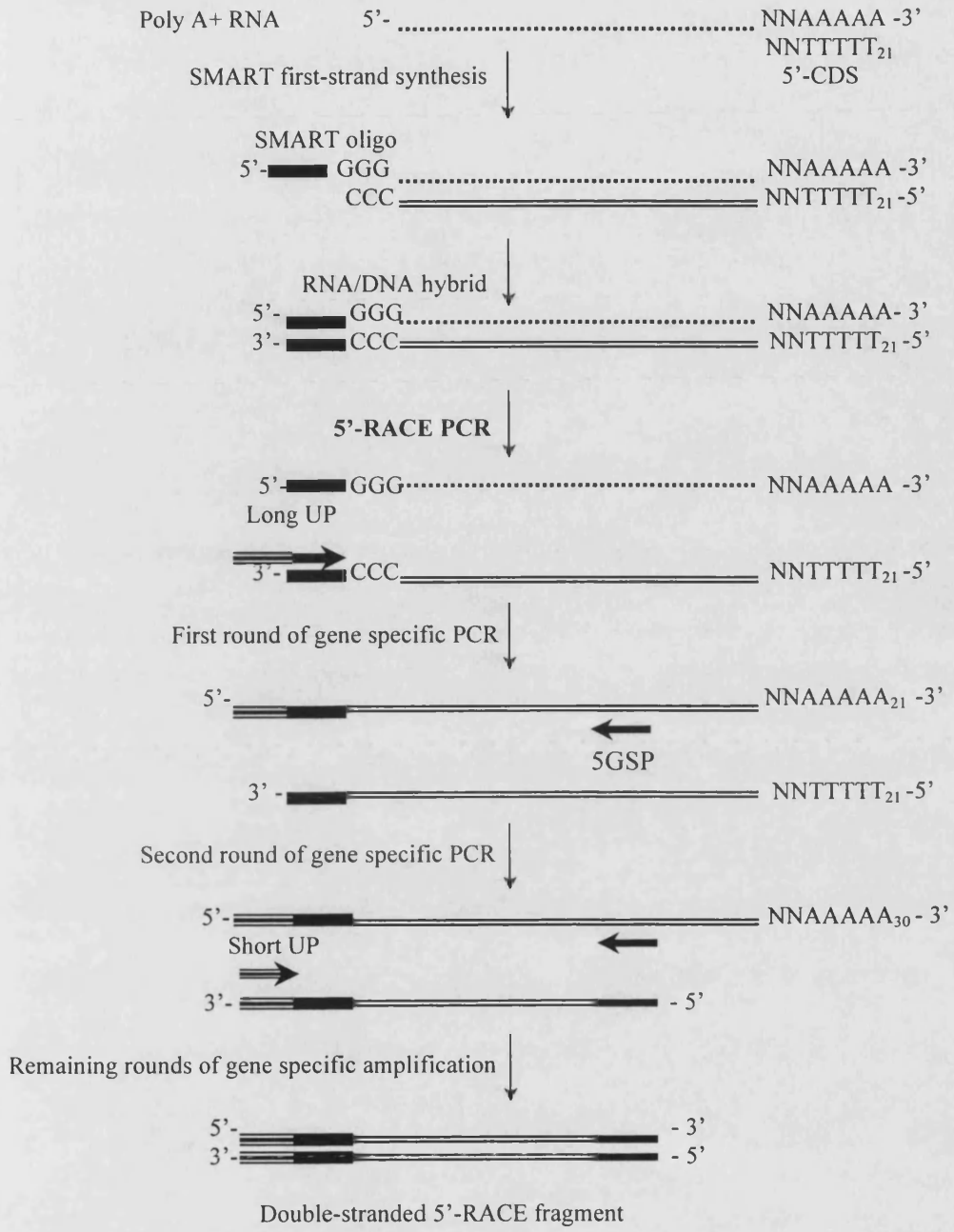



**Figure 3.7:** Flow Chart of the 5'RACE PCR reaction, modified from the Clontech SMART RACE cDNA Amplification Kit User Manual.

The strategy behind the 5'RACE PCR reaction is explained in the text. 5GSP, 5'gene specific primer; Long UP, Long universal primer; Short UP, Short universal primer

Gene Specific Primer (5GSP).

The strategy to clone a 3'RACE product is based on a similar principle (Figure 3.8). The 3'RACE cDNA population is synthesised by reverse transcribing mRNA using the 3' cDNA Synthesis primer (3'CDS). It represents an oligo(dT) primer with the SMART oligo sequence linked to its 5'end. During the first round of gene specific 3'RACE PCR, a fragment is amplified between the 3' Gene Specific Primer (3GSP) and a Long UP, which recognises the SMART oligo sequence. Again, this enables the isolation of a full length 3'RACE clone by extending the mRNA sequence past the polyA-tail. During subsequent rounds of gene specific synthesis, amplification occurs between the 3GSP and a Short UP, which anneals to the 5' end of the Long UP.

The SMART RACE PCR Kit contains all primers apart from the 5' and 3' Gene Specific Primers (5GSP and 3GSP). Thus, two gene specific and two nested gene specific primers located within the *SoxS4* HMG box (Figure 3.9) were designed with the help of the primer design program Netprimer. Parameters in the program were adjusted according to guidelines in the Clontech RACE PCR manual, which included length of primer, melting temperature Tm, %GC content and number of GC clamps at the 3'end. %GC content is the percentage of guanines (dG) and cytosines (dC) in the primer sequence. Adenine forms 2 hydrogen bonds with thymine, whereas guanine forms 3 hydrogen bonds with cytosine. Hence, a high %GC content increases the stability of the primer bound to the template and thus raises the Tm of the primer. The rationale for designing primers with at least 1 GC pair at the 3'end is based on the same principle. A GC clamp increases the strength of bonding of the primer to the DNA template at its 3'end. This is important since the DNA polymerase can only synthesise the complementary strand once the primer's 3'end has annealed.

The sequence of the primers can be found in the Materials and Methods and location of the primers is shown in Figure 3.9.

Control reactions supplied in the kit were set up to optimise the PCR program for our samples and PCR machine. They involved synthesis of control 5' and 3'RACE cDNA populations and amplification of control products using supplied gene specific primers. Bands of the expected size were amplified from the control cDNAs (data not shown), but products were faint. As a result, the cycle number in the PCR program (Materials and Methods) was increased by 5 cycles as suggested in the kit's guidelines.

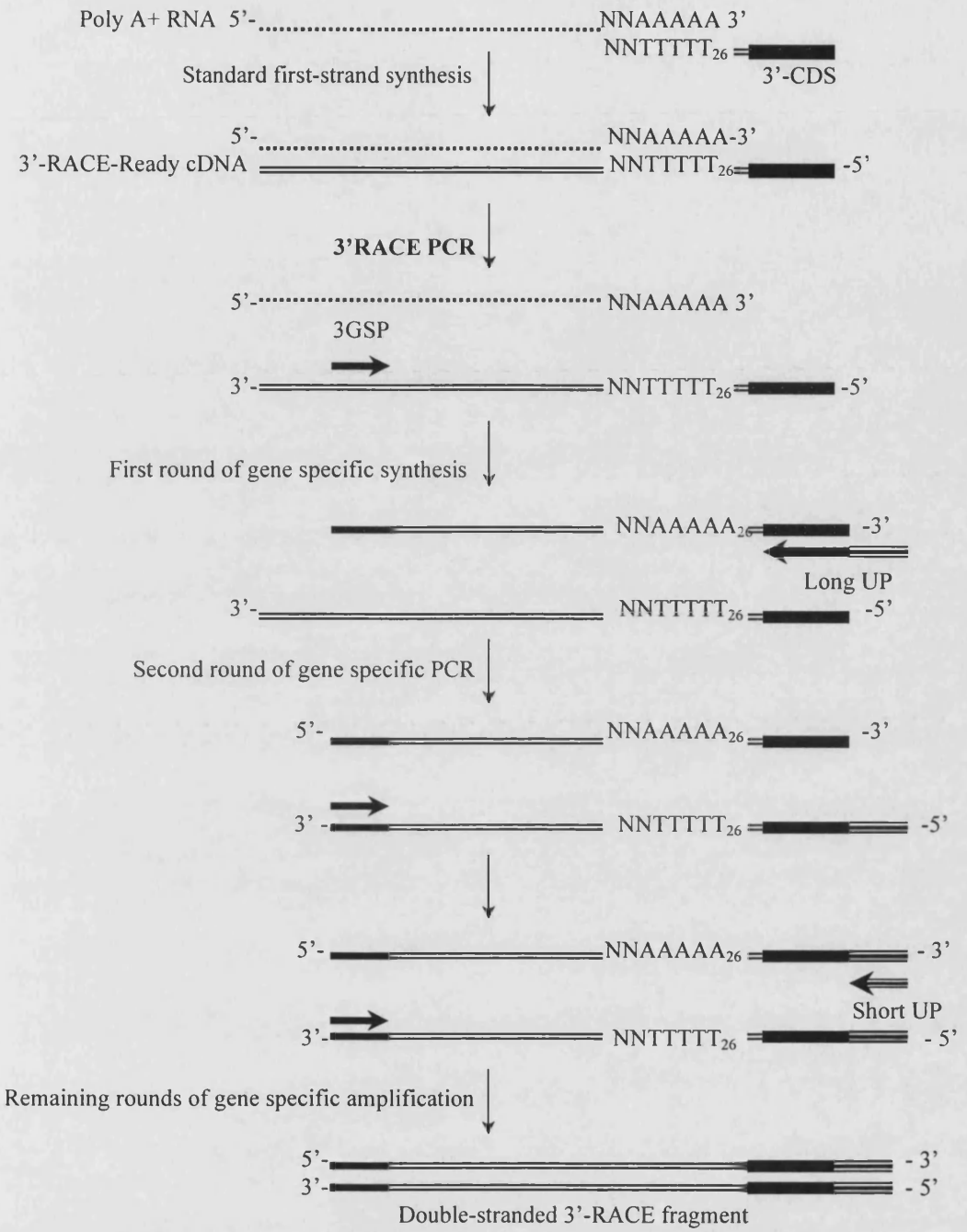



**Figure 3.8:** Flow Chart of the 3'RACE PCR reaction, modified from the Clontech SMART RACE cDNA Amplification Kit User Manual.

The strategy behind the 3'RACE PCR reaction is explained in the text. 3' GSP, 3' gene specific primer; Long UP, Long universal primer; Short UP, Short universal primer

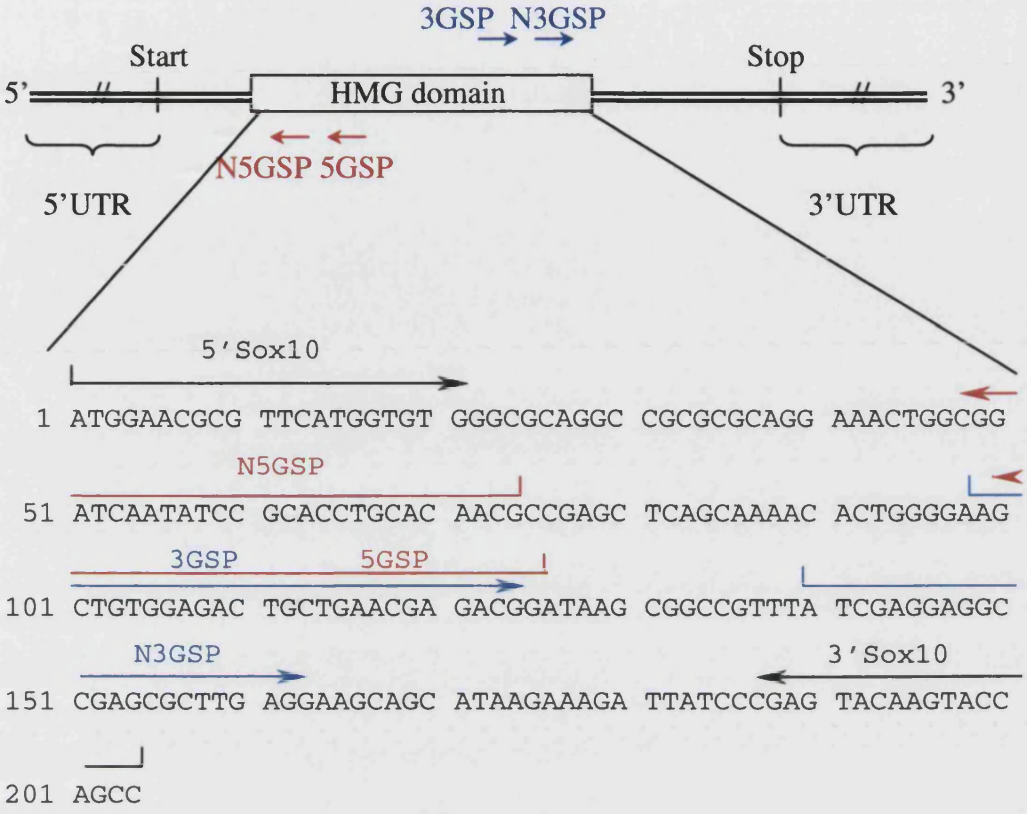



**Figure 3.9:** Location of gene specific primers used in the RACE PCR.

The figure shows the general structure of a Sox10 homologue including the DNA binding domain (HMG domain). Start and Stop mark the ends of the coding region, which is preceeded by a 5'untranslated region (5'UTR) and followed by a 3'untranslated region (3'UTR). The degenerate primers 5'Sox10 and 3'Sox10 in black were used to amplify the HMG box clone *SoxS4* by RT-PCR. The 5' gene specific primer (5GSP) and the nested 5' gene specific primer (N5GSP) in red located towards the N-terminal of the HMG box clone *SoxS4* were used to amplify a 5'RACE fragment. Similarly, the 3' gene specific primer (3GSP) and nested 3' gene specific primer (N3GSP) in blue towards the C-terminal end of *SoxS4* were used to amplify a 3'RACE product.

### *Cloning of the region 3' to the HMG box by 3'RACE PCR*

A schematic of the 3'RACE PCR Flow Chart is depicted in Figure 3.8.

The same total RNA sample as for the RT-PCR of the HMG box fragment was used to synthesise the 3' RACE cDNA population. The integrity of the 19hpf stage total RNA was confirmed by gel electrophoresis beforehand (data not shown).

PCR and single primer control reactions were set up and examined by gel electrophoresis. Two very faint bands of approximately 1.2kb and 500bp length were observed, but none in the single primer control reactions (data not shown).

To test, whether these bands were specific products or just PCR artefacts, a nested PCR was set up with 1:25 diluted 3'RACE fragment. The nested PCR protocol amplified a product of approximately 1.8kb as expected (Figure 3.10) between the nested gene specific primer (N3GSP, Figure 3.9) and the universal primer mix (UPM) containing the Short and the Long UP. Bands of 1.2kb and 500bp were also observed again, but were thought to be too short to be full length products based on comparisons of human and mouse *SOX10/Sox10* sequences (Pingault et al., 1998a; Pusch et al., 1998).

This nested 3'PCR product was directly cloned into the pGEM-T vector and plasmids prepared from white colonies. *Pvu*II digests were used to screen for clones with the expected insert of 2.25kb. Of 24 clones tested, 2 (clone B2 and C2) showed the correct sized insert (Figure 3.10). A full sequence was obtained using the automated sequencing facility provided in the Department of Biology and Biochemistry. The sequence of clones B2 and C2 were identical. A BLASTN search revealed homologies of the 3'RACE clones to *Sox10* genes. However, surprisingly, they aligned to a different region within the mouse *Sox10* sequence than expected. Thus, the forward primer (N3GSP) had bound to a similar, but not identical, site approximately 1000bp further downstream in the corresponding mouse sequence (Figure 3.11). This finding also indicated that the 3'UTR (3' untranslated region) in zebrafish was approximately 1000bp larger than in mice or humans. As expected, no significant sequence homology was observed in 3'UTRs between mammalian and zebrafish *sox10*.

New gene specific reverse primers were designed within the 3'UTR (S9 and S13, Figure 3.11) to clone the missing fragment linking the HMG box to the 3'RACE clone (B2 or C2). Fresh RT reactions were set up using 19hpf total RNA and either random hexamer primers or oligo(dT). PCR reactions with different combinations of primers and both cDNA populations were investigated including 5'Sox10-S9, 5'Sox10-S10, 5'Sox10-S13, 3GSP-S9 and N3GSP-S9 (Materials and Methods). Generally, RT-

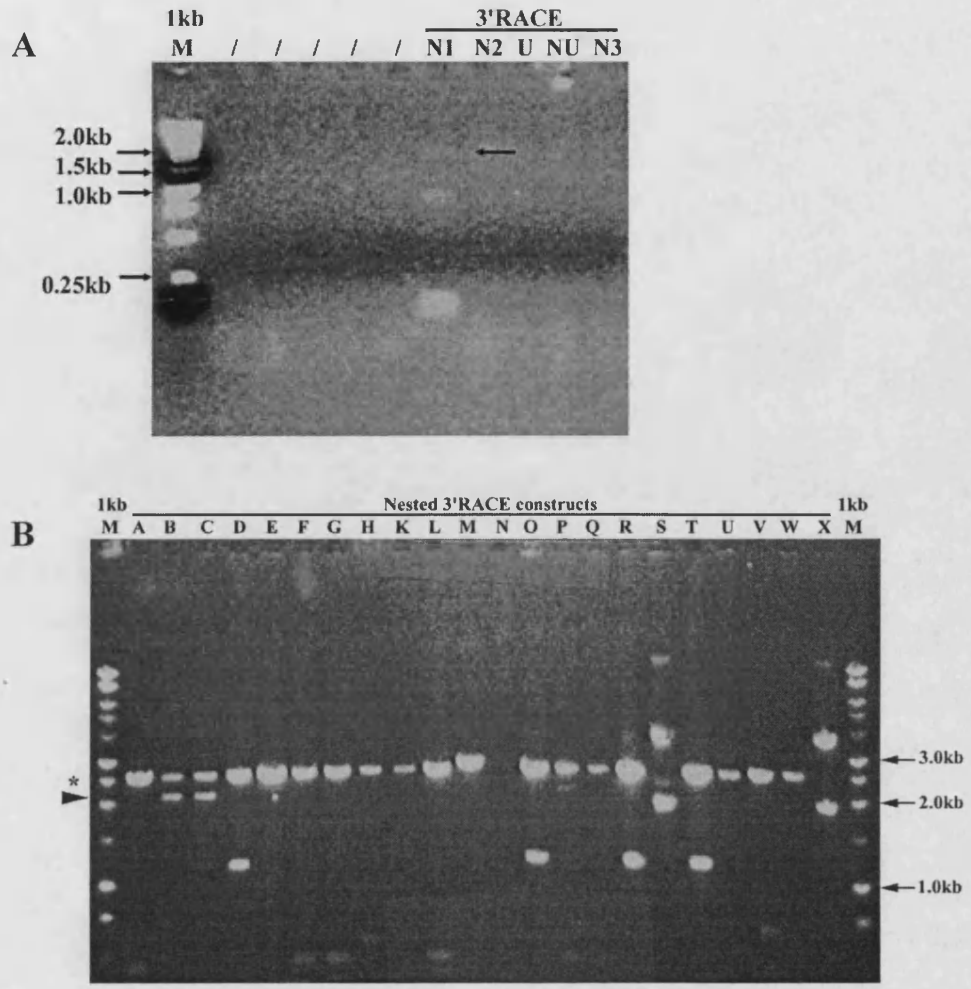
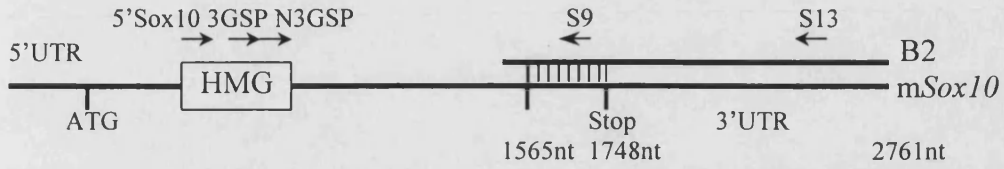



**Figure 3.10:** Cloning of the 3' RACE product

(A) Agarose gel electrophoresis of nested 3'RACE PCR products. Nested PCR amplification was carried out between primers N3GSP (nested 3' gene specific primer) and UPM (universal primer mix) (lane N1) and between primers N3GSP and NUPM (nested universal primer mix) (lane N2). In lane N1, a product of the expected 1.8kb size (arrow) was observed additionally to two non-specific products. The size was judged by comparison to a 1kb marker (1kb M, Promega). No such band was observed in lane N2. Single primer control reactions were set up with UPM only (lane U), NUPM only (lane NU) and N3GSP only (lane N3). As expected, no products were obtained in these control reactions. (B) Samples of plasmid DNA minipreparation were digested with *Pvu*II to test for inserts. A positive clone was expected to show bands of 2.25kb (insert, arrowhead) and 3kb (linearised vector, asterisk). Two positives (lane B and C) out of a total of 24 digests were identified. Sizes were estimated by comparison to 1kb markers (1kb M, Promega).

**A**




**B**

**alignment**

```
B2     : 80   caggcctcgggatcctactacgcccactccagccagacctcaggcctctactccgccttc 139
                ||| |||| ||| |||| || | |||  |  ||||| |||| |||||||| || |||||
mSox10: 1565 cagccctcaggaccctattatggccatgcaggccaggcctctggcctctattcagccttt 1624

B2     : 140  tcctacatgggcccctcacagcggcccctgtacaccgccattccggatccggg---atcc 196
                || |||||||| ||||| |||||||||||| ||||| |||||  | || ||  |    ||
mSox10: 1625 tcttacatggggccctcccagcggcccctctacactgccatctctgaccccagcccctca 1684

B2     : 197  gtgccgcagtcacacagccctacgcattgggagcagcccgtatacaccacactgtctcga 256
                | ||| || || ||||| || || || ||||||||||| ||||| || || || || |||
mSox10: 1685 gggccccaatcccacagtcccacacactgggagcagccagtatatacgactctatcccga 1744

B2     : 257  cc 258
                ||
mSox10: 1745 cc 1746
```

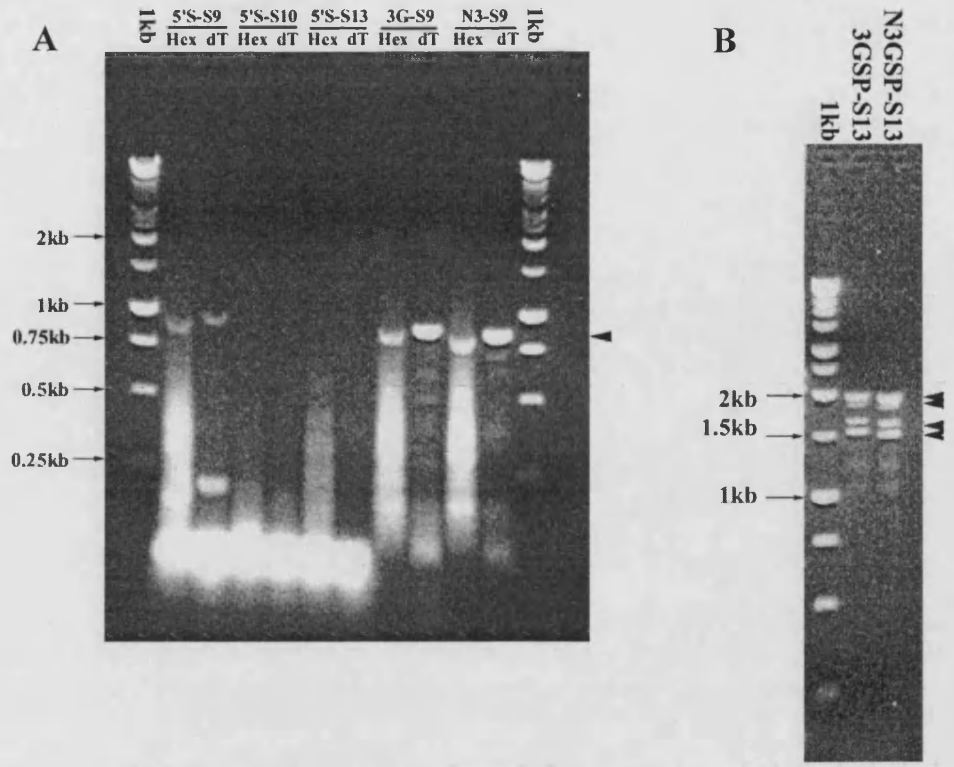
**Figure 3.11:** BLAST alignment of clone B2 and mouse *Sox10*

(A) The schematic shows mouse *Sox10* and the region of high homology to the 3'RACE PCR clone B2 (dashed box) corresponding to the C-terminus of the coding region. The B2 3'RACE clone did not overlap with the HMG box clone SoxS4. To clone the missing fragment, PCR amplifications were set up between primers 5'Sox10, 3GSP (3' gene specific primer) and N3GSP (nested 3' gene specific primer) located within the HMG box and newly designed gene specific reverse primers S9 and S13. Numbers refer to nucleotide positions in the mouse *Sox10* gene. 5'UTR, 5' untranslated region; 3'UTR, 3' untranslated region. (B) Alignment of the homologous region ( ▥ ) between 3'RACE clone B2 and mouse *Sox10* created with BLAST (Alignment of two Sequences, Materials and Methods). The region extends between nucleotide positions in mouse *Sox10* 1565 and 1746, close to the stop codon at position 1748. This corresponds to the C-terminal coding region and, as anticipated, no significant identity was observed in the 3'UTR. Note that although the 3'RACE reaction was meant to amplify between 3GSP and the 3'CDS (an oligo (dT) supplied in the SMART RACE Kit, Figure 3.8), the 3GSP must have bound further downstream and produced an approximately 1000bp shorter product.

PCR reactions from oligo(dT) cDNA populations produced fewer non-specific products and curiously the bands derived were consistently slightly larger than equivalent products amplified from random hexamer derived cDNA (Figure 3.12A). Reactions with 3GSP and N3GSP primers in combination with S9 produced a strong band of approximately 800bp. This was consistent with the size expected for this fragment from the equivalent *Sox10* region in mouse or human (Pingault et al., 1998a; Pusch et al., 1998). Furthermore, as expected, the 5'Sox10-S9 product was approximately 100bp longer. This corresponds to the distance between the 5'GSP and the N3GSP primer within the HMG box. However, in all those reactions non-specific products were abundant. A second round of amplification was used to bias the reaction towards amplifying a specific product. Thus, nested PCR reactions were set up using the first round PCR reaction that contained the least number of non-specific products, 5'Sox10S13 (oligodT), and either 3GSP or N3GSP and S13 as nested primers. The same PCR program was used, but the annealing temperature was raised to 66°C. The size of the expected product was estimated to be approximately 1.5-2kb. When PCR reaction were analysed by gel electrophoresis, 3-4 products ranging between 1.5 and 2.0kb were observed (Figure 3.12B). Since all 4 were equally strong products it was hypothesised that they might all represent specific products, maybe even different splicing variants. The entire reactions N3GSP-S9, 3GSP-S9, N3GSP-S13 and 3GSP-S13 were shotgun cloned into pGEM-T Easy vector. Five white colonies from each of the 4 transformations were chosen using criteria described in section 3.1. Plasmids were prepared using the Wizard Plus SV Miniprep DNA Purification System (Promega). After digestion with *Eco*RI, 3 constructs, 15, 19 and 20, contained inserts of approximately the correct size. Only clones 15 and 20 were sent off for automated sequencing (Figure 3.12C). Apart from a 109bp insertion in the 3'UTR of clone 15, which is not present in clone 20, they were identical to each other. Both clones showed high homology within the coding region to other *Sox10* genes (for example, 88 out of 104 nucleotides identical to chicken *Sox10*). Only 51 out of 60 nucleotides were identical to the best matching *Sox9* homologue from chicken. No homologies were observed to *Sox8* homologues in a BLAST search.

With the help of the gelassemble program, clones SoxS4, 20 and B2 were assembled into a single contig (Figure 3.13).

**Figure 3.12:** Cloning of a fragment linking the HMG box region to the 3'RACE clone (A) RT-PCRs with various primer combinations were tested by agarose gel electrophoresis for amplification of expected-sized fragments. Generally, reactions from oligo(dT) derived cDNA ("dT" lanes) gave cleaner products with less non-specific smear with all primer combinations than reactions from random hexamer derived cDNA ("Hex" lanes). Bands of approximately 800bp (arrowhead) in reactions containing primer S9 (lanes 3G-S9 and N3-S9) correspond to the estimated size of an expected product compared to a 1kb marker (1kb M, Promega). Hex, Random Hexamers; dT, oligo(dT); 5'S, 5'Sox10; 3G, 3GSP; N3, N3GSP. (B) Second round PCR amplification from the diluted 5'Sox10-S13 first round PCR reaction (lane 5'S-S13) with nested primers 3GSP/N3GSP and S13 produced 4 equally strong bands (arrowheads), which all might represent specific products. (C) Samples of plasmid DNA minipreparations were digested with *Eco*RI to test for inserts. Here, 9 such digests are shown (lanes 12-15 derived from the N3GSP-S13 PCR, lanes 16-20 derived from the 3GSP-S13 PCR), together with a 1kb marker (1kb M, Promega). A positive clone was expected to show bands of 1.5-2.0kb (insert, arrowhead) and 3kb (linearised vector, asterisk). Clones 15, 19 and 20 containing inserts of the expected size, but only clones 15 and 20 were sequenced.


A
1kb
5'S-S9 Hex dT
5'S-S10 Hex dT
5'S-S13 Hex dT
3G-S9 Hex dT
N3-S9 Hex dT
1kb
2kb
1kb
0.75kb
0.5kb
0.25kb
B
N3GSP-S13
3GSP-S13
1kb
2kb
1.5kb
1kb
C
1kb M
N3GSP-S13
12 13 14 15
3GSP-S13
16 17 18 19 20
λ-HindIII
3kb
2kb
1kb
0.75kb
*

### *Cloning of the region 5' to the HMG box by 5'RACE PCR*

A separate 5'RACE cDNA population was synthesised according to the manufacturer's instructions (see Materials and Methods). This was necessary since the SMART oligo sequence was not incorporated into the 3'CDS primer (Figure 3.8), but used as a separate primer in the 5'RACE PCR extending the 5'end of the cDNA (Figure 3.7). RACE PCR reactions were set up amplifying the region between the SMART oligo supplied in the kit and the 5'gene specific primer (5GSP). Initial attempts failed to produce any bands. There are several possible reasons that could explain such a result. For example, the 5GSP primer might be poorly designed. The latter might cause mispriming to several locations on the template, to the SMART oligo primer or to the Long Universal Primer. However, it seemed more likely that the quality of the 5'RACE cDNA was insufficient. This could be due to difficulties of the reverse transcriptase to synthesise full length transcripts in the presence of high secondary RNA structures such as hairpins or GC rich regions.

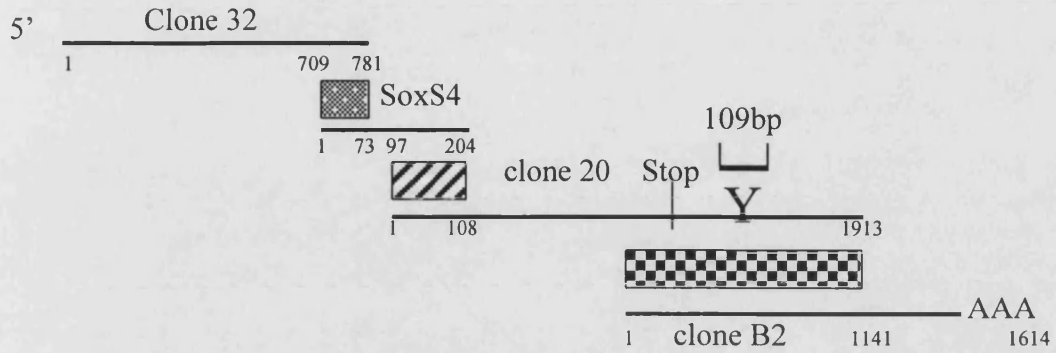
Thus, it was decided to resynthesise the 5'RACE cDNA ("new") with standard random hexamer primers instead of the 5'CDS ("old"). Both cDNA populations were then tested in RT-PCR reactions, whether RNAseH treatment and/or addition of dimethylsulphoxide (DMSO) to the PCR reaction would improve the amplification. RNAseH digests RNA in RNA/DNA hybrids and thus might improve primer binding conditions. Addition of DMSO might lead to a decrease in cDNA secondary structures and hence facilitate primer access to binding sites.

A product of the expected size was obtained from both "old" and "new" cDNAs when treated with RNAseH alone (see Figure 3.14A). Surprisingly, a similar product was obtained from the "old" cDNA population that had not been treated with either RNAseH or DMSO. To avoid non-specific by-products, nested PCR reactions were set up with 1:25 diluted first round PCR products and primers N5GSP and UPM. Agarose gel electrophoresis of samples showed clean products of approximately 850bp and 800bp (Figure 3.14B). PCR products derived from one of each cDNA reaction, "-R-D, o" and "+R-D, n" were cloned into pGEM-T Easy vector. 5 out of 10 plasmid minipreparations tested contained an *Eco*RI fragment of the correct size (Figure 3.14C). Two constructs (32 and 62) were sequenced and their sequences matched perfectly, although clone 32 had a 26bp 5'extension compared with clone 62. When assessed in a BLAST search, clones 32 and 62 showed high identity to regions of human *SOX10* (161/183bp identical) and human *SOX9* (160/183bp identical) corresponding to the

**Figure 3.13:** Schematic diagram showing the overlap of *sox10*-like fragments of clones 32, SoxS4, 20 and B2

(A) Schematic diagram of overlapping fragments. Numbering refers to nucleotide positions in individual clones. Clone 15 and 20 were identical apart from a 109bp insert in clone 15. The site of insertion in clone 20 is indicated. (B) The three regions of overlap are shown. Clone 32 overlaps with clone SoxS4 over a region of 73bp in the N-terminal region of the HMG box ( ▩ ), whereas clone SoxS4 overlaps with clone 20 in the 108bp C-terminal region of the HMG box ( ▨ ). Clones 20 and B2 align completely over a 1141bp region ( ▣ , A), but only the first 120bp and the last 121bp are shown in (B).

**A**




**B**

 alignment

```
32   : 709 atgaacgcgttcatggtgtgggcgcaggccgcgcgcaggaaactggcggatcaatatccg 768
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
SoxS4: 1   atgaacgcgttcatggtgtgggcgcaggccgcgcgcaggaaactggcggatcaatatccg 60

32   : 769 cacctgcacaacg 781
        |||||||||||||
SoxS4: 61  cacctgcacaacg 73
```

 alignment

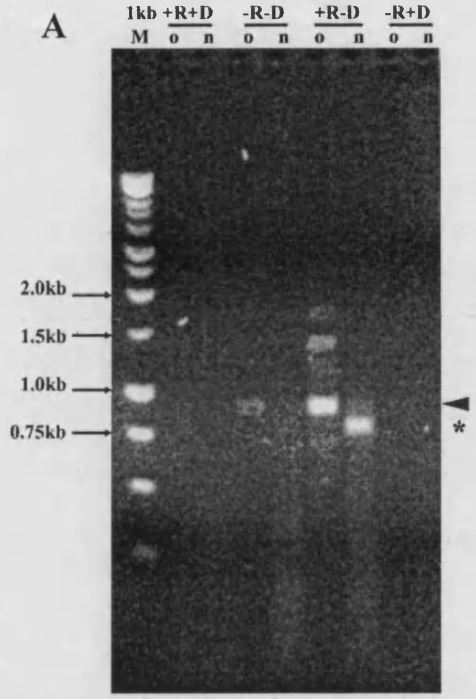
```
20   : 1   aagctgtggagactgctgaacgagacggataagcggccgtttatcgaggaggccgagcgc 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
SoxS4: 97  aagctgtggagactgctgaacgagacggataagcggccgtttatcgaggaggccgagcgc 156

20   : 61  ttgaggaagcagcataagaaagattatcccgagtacaagtaccagcca 108
        |||||||||||||||||||||||||||||||||||||||||||||||||||
SoxS4: 157 ttgaggaagcagcataagaaagattatcccgagtacaagtaccagcca 204
```

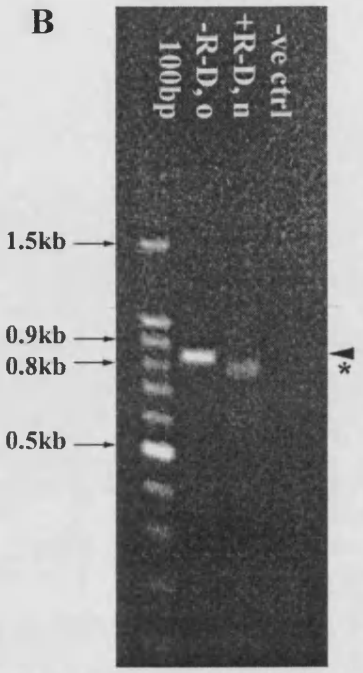
 alignment

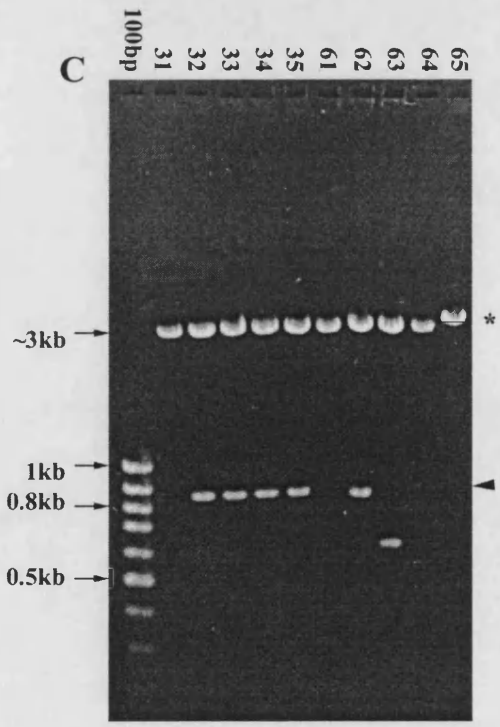
```
B2   : 1    tcccgcgcacaattcgccgaatacgccgagcaccaggcctcgggatcctactacgcccac 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20   : 773  tcccgcgcacaattcgccgaatacgccgagcaccaggcctcgggatcctactacgcccac 832

B2   : 61   tccagccagacctcaggcctctactccgccttctcctacatgggcccctcacagcggccc 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20   : 833  tccagccagacctcaggcctctactccgccttctcctacatgggcccctcacagcggccc 892
-------------------------------------------------------------------------------
B2   : 1021 ataaatgactatttttagagaataccgctagtgcctcaagtccatcacaaacgaattgtc 1080
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20   : 1793 ataaatgactatttttagagaataccgctagtgcctcaagtccatcacaaacgaattgtc 1852

B2   : 1081 gtttctgataaattcaattttgatgatgtaaaatccttttcagggtgcaaattttatcat 1140
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
20   : 1853 gtttctgataaattcaattttgatgatgtaaaatccttttcagggtgcaaattttatcat 1912

B2   : 1141 g 1141
        |
20   : 1913 g 1913
```

**Figure 3.14:** Cloning of the 5'RACE PCR fragment

(A) Multiple RT-PCR reaction conditions were evaluated for amplification of expected 5' sox10 fragments. RT-PCR were set up with 5'RACE cDNA (old, o) and cDNA made with Random Hexamer primers (new, n), with or without RNAseH (+R, -R) and with or without DMSO (+D, -D) treatment. A band of approximately the expected size (900bp) was observed in reaction lanes "-R-D, o" and "+R-D, o" (arrowhead), whereas in lane "+R-D, n" the product appeared to be slightly smaller (asterisk) compared to a 1kb marker (1kb M, Promega). Two non-specific products were also observed in reaction "+R-D, o". (B) Second round PCR amplification from "-R-D, o" and "+R-D, n" with nested primers N5GSP (5' nested gene specific primer) and UPM (universal primer mix) obtained single products of the correct size. The nested product of reaction "-R-D, o" (arrowhead) was approximately 850bp, the band from "+R-D, n" approximately 800bp (asterisk). As expected, no product was obtained in the control reaction without a DNA template (lane "-ve ctrl"). A 100bp marker (100bp, Promega) was used to judge sizes. A nested reaction from reaction "+R-D, o" was omitted due to presence of strong non-specific products. (C) Samples of plasmid DNA minipreparations were digested with *Eco*RI to test for inserts. Here, 10 such digests are shown (lanes 31-35 derived from "-R-D, o" PCR and lanes 61-65 derived from "+R-D, n" PCR), together with a 100bp marker (100bp, Promega). Positive clones were expected to show bands of approximately 850bp (inserts, arrowhead) and 3kb (linearised vector, asterisk). Clones 32, 33, 34, 35 and 62 were identified as likely positives and clones 32 and 62 were sequenced.


A
1kb
M
+R+D
o
n
-R-D
o
n
+R-D
o
n
-R+D
o
n
2.0kb
1.5kb
1.0kb
0.75kb
*


B
100bp
-R-D, o
+R-D, n
-ve ctrl
1.5kb
0.9kb
0.8kb
0.5kb
*


C
100bp
31
32
33
34
35
61
62
63
64
65
~3kb
1kb
0.8kb
0.5kb
*

HMG box region and the sequence just upstream. In comparison, only 79/91bp were identical to human *SOX8*.

The sequence of the slightly longer 32 clone was used to assemble a consensus sequence of the full length cDNA (Figure 3.13A).

### *Assembly of a full length* sox10*-like cDNA*

All sequences were assembled into a single 3231bp contig (Figure 3.13A) and the consensus sequence, called *sox10*-like, deduced using the GCG package. The sequence of the 109bp insertion of clone 15 in the 3'UTR was included in the consensus sequence since its presence had been confirmed in the genomic sequence (T. Carney, pers. comm.). All sequence ambiguities were solved by sequencing both strands of all clones, 32, 62, S4, 15, 20, B2 and C2 several times and where necessary, using internal primers to resolve persistently ambiguous bases. As will be described in chapter 6, the wild-type coding region, 1455bp (485 residues) in length, was additionally confirmed by sequencing 4 overlapping PCR fragments multiple times in the search for mutant lesions. The consensus sequence submitted to Genbank (accession number AF402677) is shown in Figure 3.15. A 5'UTR of 375bp preceeds the predicted open reading frame which is followed by a 3'UTR of 1398bp.

## Sequence alignment with other *sox10* homologues and phylogenetic analysis

To assign a definite identity to the *sox10-like* gene we used BLAST searches, multiple sequence alignments with closest homologues of the same subgroup and phylogenetic analysis.

Firstly, a BLASTX search with the complete predicted protein sequence reveals high amino acid identities to mouse Sox10 protein (266/435 residues) and less to mouse Sox9 (217/466 residues) and mouse Sox8 (192/437 residues).

ClustalW alignment of the full length predicted proteins and calculation of a sequence identity matrix (BioEdit program) shows that the isolated Sox10-like clone shows 61% and 60% amino acid identity to mammalian and chicken Sox10, respectively, but only about 40% identity to zebrafish Sox9 paralogues, mammalian Sox9 and trout and mammalian Sox8 homologues (Table 3.3).

Secondly, the Sox10-like predicted open reading frame appears to have all domains

**Figure 3.15:** The cDNA sequence of a zebrafish *sox10-like* homologue as published in Genbank, accession number AF402677

The predicted 485 residue coding region extends between nucleotides 376 and 1833. The highly conserved DNA binding domain (HMG box) is found between nucleotides 691 and 930. These 80 residues are marked in red. A 375bp 5' untranslated region (5'UTR) is located N-terminal to the start of translation between nucleotides 1 and 375, whereas a 1398bp 3' untranslated region (3'UTR) at the C-terminus is found between 1834-3231. Although no attempt was made to investigate whether the cDNA was full length at the N-terminal end, the C-terminal end appears complete with a 28bp polyA-tail.

```
   1  GTTTTTTCTC GCTCTTCACA CAACGGGGCT CTTTAAGCCT CGACGCGCGA CACAGAGCAG GCATTCAGAG
  71  CGCGAGCGAG GGGGCTGAAC CGACGGACTC TCGCGCTGGG CGGGCGACCT GCTGCACTGT AAAGTTTCCA
 141  TCAGATCTAT ATTCTGAGGA AGACGGCGGA AGGATTCCTT CTGACAGAGT CGAGTCGTTT GAAAAGAACT
 211  GTTAAAGTTT CACTGGATGA TCTTAAATAA TAAACAAAAG CACAATTATT TTACAAGAAA AAAACATTTG
 281  AGAAGTATAA ATTAATACAT TTATATTTAA AATAAAATTT AAGTGAGGAA ATTAAACCTA CCGAAGTCAC
 351  CTGTGGCCGC AGAACTAGTG GACCGATGTC GGCGGAGGAG CACAGCATGT CGGAGGTGGA AATGAGTCCC
   1                          M   S  A  E  E   H  S  M   S  E  V   E  M  S  P
 421  GGGGTGTCGG ACGATGGGCA CTCCATGTCC CCTGGTCACT CGTCGGGCGC TCCCGGTGGC GCGGACTCCC
  16    G  V  S   D  D  G   H  S  M  S   P  G  H   S  S  G   A  P  G  G   A  D  S
 491  CTCTGCCCGG TCAGCAGTCT CAGATGTCCG GGATCGGGGA TGATGGAGCC GGTGTCTCCG GCGGGGTCTC
  39   P  L  P   G  Q  Q  S   Q  M  S   G  I  G   D  D  G  A   G  V  S   G  G  V
 561  GGTGAAGTCC GACGAGGAAG ATGACCGGTT CCCCATCGGC ATCCGCGAGG CGGTCAGTCA GGTGCTGAAC
  62 S  V  K  S   D  E  E   D  D  R   F  P  I  G   I  R  E   A  V  S   Q  V  L  N
 631  GGGTACGACT GGACGCTCGT GCCCATGCCC GTGCGCGTGA ACTCGGGCAG CAAGAGCAAA CCGCACGTCA
  86   G  Y  D   W  T  L   V  P  M  P   V  R  V   N  S  G   S  K  S  K   P  H  V
 701  AGCGGCCGAT GAACGCGTTC ATGGTGTGGG CGCAGGCCGC GCGCAGGAAA CTGGCGGATC AATATCCGCA
 109  K  R  P   M  N  A  F   M  V  W   A  Q  A   A  R  R  K   L  A  D   Q  Y  P
 771  CCTGCACAAC GCCGAGCTCA GCAAAACACT GGGGAAGCTG TGGAGACTGC TGAACGAGAC GGATAAGCGG
 132 H  L  H  N   A  E  L   S  K  T   L  G  K  L   W  R  L   L  N  E   T  D  K  R
 841  CCGTTTATCG AGGAGGCCGA GCGCTTGAGG AAGCAGCATA AGAAAGATTA TCCCGAGTAC AAGTACCAGC
 156   P  F  I   E  E  A   E  R  L  R   K  Q  H   K  K  D   Y  P  E  Y   K  Y  Q
 911  CACGTCGACG CAAGAACGGC AAACCGGGTT CCAGCTCAGA GGCCGACGCC CACTCTGAGG GTGAGGTCAG
 179  P  R  R   R  K  N  G   K  P  G   S  S  S   E  A  D  A   H  S  E   G  E  V
 981  CCACAGCCAA TCGCATTACA AGAGCCTGCA CCTGGAGGTG GCGCACGGCG GGGCTGCAGG GTCACCATTG
 202 S  H  S  Q   S  H  Y   K  S  L   H  L  E  V   A  H  G   G  A  A   G  S  P  L
1051  GGTGATGGAC ACCACCCTCA CGCTACAGGT CAGAGTCACA GCCCTCCAAC GCCCCCTACC ACCCCCAAGA
 226   G  D  G   H  H  P   H  A  T  G   Q  S  H   S  P  P   T  P  P  T   T  P  K
1121  CGGAACTGCA GGGAGGAAAA TCAGGCGAGG GCAAGCGTGA GGGCGGAGCC TCTCGGAGTG GACTGGGGGT
 249   T  E  L   Q  G  G  K   S  G  E   G  K  R   E  G  G  A   S  R  S   G  L  G
1191  GGGAGCAGAT GGAAGCTCCG CCTCATCGTC TGCCAGCGGG AAACCGCACA TCGACTTCGG TAACGTGGAC
 272 V  G  A  D   G  S  S   A  S  S   S  A  S  G   K  P  H   I  D  F   G  N  V  D
1261  ATTGGCGAAA TCAGCCATGA CGTGATGGCC AACATGGAGC CGTTCGACGT GAACGAGTTC GACCAGTATC
 296   I  G  E   I  S  H   D  V  M  A   N  M  E   P  F  D   V  N  E  F   D  Q  Y
1331  TCCCACCCAA TGGCCACCCG CAGGCGTCCG CCACTGCCAG CGCAGGATCT GCAGCGCCAT CGTATACATA
 319  L  P  P   N  G  H  P   Q  A  S   A  T  A   S  A  S  G   A  A  P   S  Y  T
1401  CGGCATCTCC AGCGCGCTAG CGGCCGCTAG TGGCCACTCC ACCGCATGGC TGTCCAAGCA GCAACTGCCG
 342 Y  G  I  S   S  A  L   A  A  A   S  G  H  S   T  A  W   L  S  K   Q  Q  L  P
1471  TCCCAGCAGC ATTTGGGCGC AGATGGCGGG AAAACGCAGA TAAAGAGTGA AACACACTTC CCTGGGGATA
 366   S  Q  Q   H  L  G   A  D  G  G   K  T  Q   I  K  S   E  T  H  F   P  G  D
1541  CAGCGGCGAG CGGTTCACAC GTCACATACA CGCCGCTAAC ACTGCCGCAC TACAGCTCCG CCTTCCCCTC
 389   T  A  A   S  G  S  H   V  T  Y   T  P  L   T  L  P  H   Y  S  S   A  F  P
1611  GCTGGCGTCC CGCGCACAAT TCGCCGAATA CGCCGAGCAC CAGGCCTCGG GATCCTACTA CGCCCACTCC
 412 S  L  A  S   R  A  Q   F  A  E   Y  A  E  H   Q  A  S   G  S  Y   Y  A  H  S
1681  AGCCAGACCT CAGGCCTCTA CTCCGCCTTC TCCTACATGG GCCCCTCACA GCGGCCCCTG TACACCGCCA
 436   S  Q  T   S  G  L   Y  S  A  F   S  Y  M   G  P  S   Q  R  P  L   Y  T  A
1751  TTCCGGATCC GGGATCCGTG CCGCAGTCAC ACAGCCCTAC GCATTGGGAG CAGCCCGTAT ACACCACACT
 459   I  P  D   P  G  S  V   P  Q  S   H  S  P   T  H  W  E   Q  P  V   Y  T  T
1821  GTCTCGACCG TGACACACTC TACCAAGATG ACCAGTCACT AAAGGTCCAA CCGTAAGGTG TGTGTGTGTG
 482 L  S  R  P   *
1891  TGCTAAAAAT CATCGAAACA CTCGCCTGCA CCACAATCGA CACAAACTGA GATCTGAGAA ACGAGTGTGT
1961  GTGTGTGTGT GAGATCTGCA GGGAAATATT CTCACGTGCC TCAGACGACC ACCGTCCAGA CCTGCTCCCT
2031  CAACGCCAAT TTGACACCAG TAGTATTTTC GAAAAAGACG TAGTACCAAA GTACCGAGAC CAAAACATTA
2101  CAGAAATACG AGAGTGCATC CATCCTTCCT GAACTCCGGA TATCAGATCA CACACAGACT TCAACACATG
2171  ATGCTAGTAC CAGTGCATCC GCATTTTTTA TCTGTATTTG TATGAATGAA TAATCTTTTT ATTAACCAAA
2241  ATAAGGCCAT ATTGTTTTTA AAAAAATAAT GAGGTGTTTT TCGTTGTTGT AATCTCTTGT GTTGTCGTTA
2311  CTGTTGTTAT TTGTGTTGCC ATAACTACAC TGAAAAGTCT TCACCACTGT CTAGTGTTTG TTAATGACAT
2381  TTGTGTTTTA TGACTTTCAG CGTGTGTAAA TATCAGTGCC AGGACGCCAT ACACACATGT CTCCACCCAA
2451  TTAAGGTGCG CTCACAGTGA CGTTAATTAA ATTGAGGAAT TCCCAACCAT GCAAATTCCT CTAGAAATGG
2521  CTATATTTCG TGAAAGCAGT AAATGTGAGC GCACCTTTAT TTCACACAAG CAGTACTGTA AAGGTAATAT
2591  ATTTTTGCCA GATTGGAAAT GGTGGACGTA ATTACGAATT TTTAATAATA AATGACTATT TTTAGAGAAT
2661  ACCGCTAGTG CCTCAAGTCC ATCACAAACG AATTGTCGTT TCTGATAAAT TCAATTTTGA TGATGTAAAA
2731  TCCTTTTCAG GGTGCAAATT TTATCATGCG TTAACCGATG TGATTATACA TCGAATATGC ATATGCAAAT
2801  TAATAAGTGC CATTTTTATA ATTAAAAATA CATCAACTAT CTGAAGACCT TCCTAACATT GGGTAAGTAA
2871  ATAAATACAT TTATTTCATT GTATTATTTT TGGTAAATAC AATATTTAGC TATTCTATGT TTTGTCTCCC
2941  TTTGGTACTT TATAGTTTTG TTTTTGCCCT CTTTATTATT TAGTATTATT CAGAAACAAA CAAACTCTTT
3011  TTATATATTA CAGAATATTA TTTATATTTG TTGTTGTTTT TTTTTATCAG TAGCGTTTTA TTCTGTTTGT
3081  CGTAAACCTC TGTCGTGTGC TGTTGTGGTG GGTTAAAGTG CTGTTGTTTT TCTCTGTCGG TGTAAATAGA
3151  ACTGAGAGCA GTGACTAACT TTCCTCACTC TAAATAAAGC TGCAGTCTTT ACTAAAAAAA AAAAAAAAAA
3221  AAAAAAAAAA A
       K  K  K   K
```

characteristic for Sox10 homologues. An amino acid alignment with known Sox10 proteins and the percentage identities within distinct domains between homologues is depicted in Figure 3.16. Preliminary studies of a zebrafish Sox10 genomic clone show conservation of intron sites (T. Carney, unpubl. data).

Finally, Professor Laurence D Hurst, University of Bath, calculated a phylogenetic tree containing *Sox* gene members of subgroup E. Nucleotide sequences were translated and the coding regions aligned using ClustalX (Thompson et al., 1997) to improve the reliability of the pileup. The actual nucleotide alignment was subsequently reconstructed from the protein alignment using MRTRANS (Materials and Methods). From this sequence alignment an unrooted phylogenetic tree was constructed using PUZZLE (v 4.0.2) (Strimmer, 1996; Figure 3.17). This program calculates the probability of relatedness for each included *Sox* gene homologue by maximum likelihood. The number for each branch represents a bootstrap value, which is the percentage of times this branch is placed on the tree in this specific location (Figure 3.17). The *sox10*-like homologue clearly clustered with other *Sox10*s. The bootstrap values for each branch of *Sox10* homologues were close to the maximum of 100 indicating high support for each branch and thus a very reliable cluster.

## Discussion

We have described the successful cloning of a zebrafish *sox10* homologue.
Each of two possible strategies to clone *sox10* were begun, but screening a cDNA library was not successful in our hands. The problem was high background, despite attempting to closely mimic the conditions used by Southard-Smith and colleagues (Southard-Smith et al., 1999). There are at least two possibilities that could explain the differences in results. Although the fish species included in their zooblot was not specified in their publication, it was different to zebrafish (M. Southard-Smith, person. comm.). Secondly, the genomic blots used by SouthardSmith were commercially obtained and thus, it was impossible to mimic the exact procedures used to prepare the genomic blot. Thus, we could not confirm the band observed in fish.

Meanwhile, the second approach involving RT-PCR with generic *Sox* gene primers was successful. These degenerate primers 5'Sox10 and 3'Sox10 were located in regions very conserved in *Sox* genes. They had originally been used to amplify a *Sox2*

| Gene | csox10 | csox8 | csox9 | hsox10 | hsox8 | hsox9 | msox10 | msox8 | msox9 | rsox10 | tsox9 | tsoxp1 | **zsox10** | zsox9a | zsox9b |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| csox10 | 100 | 42.9 | 48.2 | 81.9 | 42.8 | 48.0 | 82.1 | 41.4 | 47.6 | 81.4 | 48.9 | 43.5 | **59.8** | 47.2 | 43.3 |
| csox8 | --- | 100 | 47.5 | 43.8 | 68.4 | 46.0 | 44.0 | 73.5 | 45.4 | 44.0 | 47.9 | 71.6 | **41.5** | 47.6 | 45.1 |
| csox9 | --- | --- | 100 | 46.8 | 43.3 | 81.3 | 47.1 | 45.7 | 81.0 | 47.1 | 70.4 | 46.0 | **42.8** | 68.2 | 55.0 |
| hsox10 | --- | --- | --- | 100 | 43.8 | 46.9 | 98.4 | 42.2 | 46.6 | 97.4 | 48.3 | 42.7 | **60.9** | 46.8 | 43.0 |
| hsox8 | --- | --- | --- | --- | 100 | 42.4 | 43.8 | 80.3 | 41.8 | 43.6 | 43.7 | 62.5 | **41.1** | 44.6 | 43.1 |
| hsox9 | --- | --- | --- | --- | --- | 100 | 47.3 | 43.6 | 95.8 | 47.5 | 70.2 | 44.8 | **42.7** | 68.6 | 53.9 |
| msox10 | --- | --- | --- | --- | --- | --- | 100 | 42.4 | 46.9 | 98.9 | 48.9 | 43.1 | **60.9** | 47.2 | 43.2 |
| msox8 | --- | --- | --- | --- | --- | --- | --- | 100 | 43.2 | 42.2 | 46.7 | 66.3 | **40.6** | 45.3 | 42.7 |
| msox9 | --- | --- | --- | --- | --- | --- | --- | --- | 100 | 47.1 | 69.5 | 44.0 | **42.7** | 67.9 | 52.8 |
| rsox10 | --- | --- | --- | --- | --- | --- | --- | --- | --- | 100 | 48.7 | 42.9 | **60.7** | 47.2 | 43.0 |
| tsox9 | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | 100 | 47.2 | **44.7** | 76.2 | 57.5 |
| tsoxp1 | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | 100 | **41.8** | 46.1 | 45.0 |
| **zsox10** | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | **100** | **42.6** | **40.0** |
| zsox9a | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | 100 | 59.6 |
| zsox9b | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | 100 |

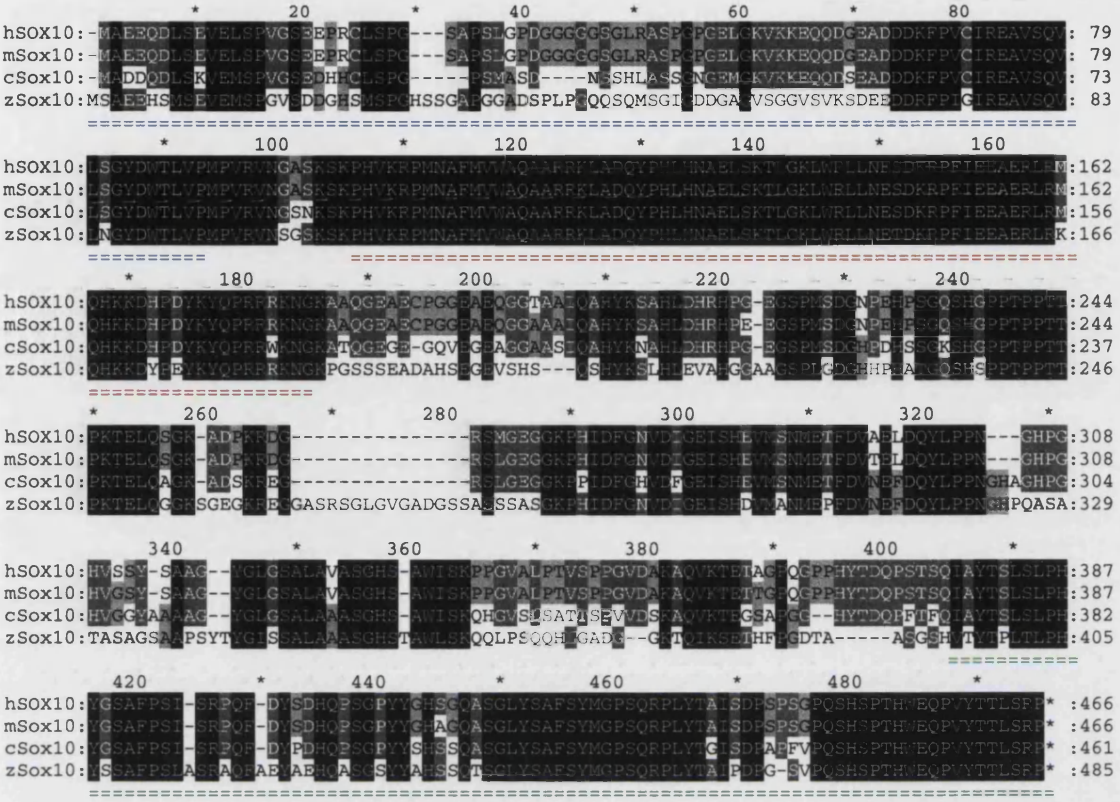
**Table 3.3:** Sequence Identity Matrix of members of subgroup E

Numbers represent % amino acid identity between two Sox genes. c, chicken; h, human; m, mouse; r, rat; t, trout; z, zebrafish.

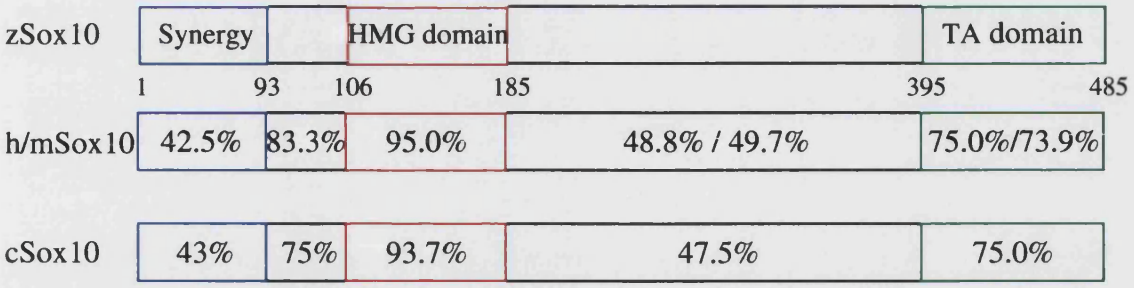
Although homologues of different species usually show 70-80% identity, zebrafish Sox10 only shows approximately 60% overall sequence identity to other Sox10s, but even less, around 40%, to Sox8s and Sox9s including both zebrafish Sox9 paralogues.

**Figure 3.16:** Amino acid identities within distinct domains between *Sox10* homologues (A) The multiple sequence alignment of published Sox10 proteins was created with Pileup (gcg package, Materials and Methods). Regions of identity are shaded in black (Genedoc). The highly conserved HMG domain is underlined red (residues 106-185 in zebrafish Sox10). The synergy domain (blue) and the transactivation (TA) domain (green) are poorly defined. However, the regions underlined in blue and green show high sequence homologies particularly between mammalian Sox10 proteins and exhibited synergistic and transactivational properties in *in vitro* studies (Kuhlbrodt et al., 1998b; Pusch et al., 1998). (B) Amino acid identities within functional regions are shown. Zebrafish Sox10 (zSox10) was compared to mammalian (hSOX10, human SOX10; mSox10, mouse Sox10) and chicken (cSox10, chicken Sox10) Sox10 proteins. The HMG domain (red) is the most conserved region with up to 95% identity. The transactivation domain (green) and the region just upstream of the HMG box with unknown function both show approximately 75% identity. However surprisingly, the synergy domain does not appear to be as highly conserved as the other domains in this zebrafish Sox10 homologue. Accession numbers of sequences are as follows: human SOX10 (hSOX10) (NM_006941); mouse Sox10 (mSox10) (AF047389); chicken Sox10 (cSox10) (AF152356); zebrafish Sox10 (zSox10) (AF402677).

**A**


hSOX10:
mSox10:
cSox10:
zSox10:

**B**


zSox10
Synergy
HMG domain
TA domain
1
93
106
185
395
485
h/mSox10
42.5%
83.3%
95.0%
48.8% / 49.7%
75.0%/73.9%
cSox10
43%
75%
93.7%
47.5%
75.0%

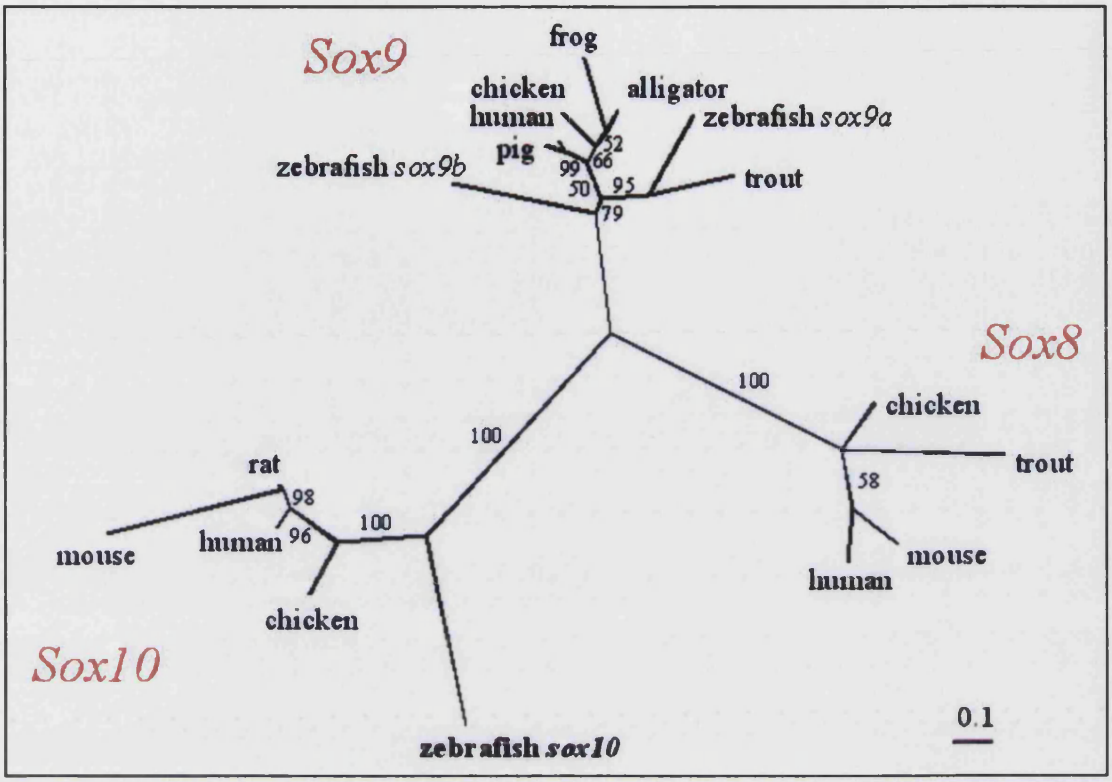



**Figure 3.17:** Maximum likelihood phylogenetic tree of subgroup E *Sox* genes generated by Prof. Laurence D Hurst, University of Bath.

Nucleotide sequences were translated and then aligned using ClustalX. The nucleotide alignments were reconstructed from the protein alignments using MRTRANS. By Maximum Likelihood an unrooted neighbour joining tree was calculated. As expected zebrafish *sox10* clearly clusters within the *Sox10* clade of vertebrate *Sox* genes. Bootstrap values close to 100 on each of the *Sox10* branches present high support for this cluster. The accession numbers for the sequences are as follows: chicken *Sox8* (AF228664); trout *SoxP1* (D83256); mouse *Sox8* (AF191325); human *Sox8* (AF226675); rana frog *Sox9a* (AB035887); alligator *Sox9* (AF106572); trout *Sox9* (AB006448); zebrafish *sox9a* (AF277096); zebrafish *sox9b* (AF277097); chicken *Sox9* (AB012236); pig *Sox9* (AF029696); human *Sox9* (Z46629); zebrafish *sox10* (AF402677); chicken *Sox10* (AF152356); mouse *Sox10* (AF047389); rat *Sox10* (AJ001029); human *Sox10* (NM_006941).

gene (Yuan et al., 1995), but also led to successful cloning of rat *Sox10* (Kuhlbrodt et al., 1998a). In the latter study, Kuhlbrodt and her colleagues used cDNA derived from a primary rat Schwann cell culture to clone Sox proteins involved in the development of Schwann cells. This cDNA might have been enriched with *Sox10* transcripts, since 69% of isolated clones showed high identity to *Sox10* homologues. Unfortunately, an equivalent source of cultured zebrafish cells was not available. Thus, to identify a *sox10* homologue, we anticipated having to screen a large number of clones. Perhaps cloning chances could have been improved by dissecting whole zebrafish embryos to obtain tissues expected to express high levels of *sox10* such as premigratory neural crest or the gut. Due to the small size of zebrafish embryos this was not attempted in this study. However, we hoped to enrich for *sox10* transcripts by including adult tissues (brain and skin). If the cloning of HMG boxes from 19hpf stage carried out in parallel had not been successful, more HMG box clones derived from adult tissues would have been screened and might have also led to success.

A total of 61 HMG box clones were manually sequenced. As expected, HMG boxes with homology to various *Sox* genes were cloned from all stages. Their identities were tentatively assigned to the closest match in a BLASTN search. This was necessary, because clones were only sequenced once and thus are prone to PCR and sequencing errors. These are the likely explanation for mismatches between the sequences obtained here and published sequences of zebrafish genes (Table 3.2).

Although the sample number was small and the cloned *Sox* gene homologues might not be representative, it was interesting that at least some could correspond to developmental processes during the stages used. Thus, 5 out of 9 HMG boxes (55.6%) cloned from adult tissue and 6 out of 10 (60%) cloned from 24hpf stage appeared to be zebrafish *sox19*. Zebrafish *sox19* was first isolated from a cDNA library prepared from late gastrulation embryos and is classified as a member of subfamily B1 (Vriz and LovellBadge, 1995). Expression studies are still unavailable, but the abundance of *sox19*-like HMG boxes observed here in adult brain might indicate a role in the CNS. It could be speculated that the high representation of *sox19*-like clones at 24hpf is also derived from expression in the developing CNS.

Two *sox11*-like homologues (subfamily C) were cloned from the 24hpf stage, one of which clearly matched the *sox11a* paralogue. This is consistent with expression of *sox11a* in the fore-, mid- and hindbrain at 24hpf (Rimini et al., 1999). One of 9 clones isolated from the 19hpf stage, 3 out of 6 clones from the 5dpf and 1 clone out of 9 from

the adult stage showed high homologies to zebrafish *sox11b*. The latter result is surprising, since Rimini reported that expression levels of *sox11b* at 36hpf were already fainter and from 48hpf no expression could be detected. This does not exclude that *sox11b* might be required again later and thus, our result might point to an unknown late function of *sox11b*.

Another two *Sox* genes, similar to *sox31* and *sox2*, also involved in the development of the CNS, were isolated. Zebrafish *sox31* was identified from the 24hpf stage. Consistent with our observations, it is expressed in the developing CNS (forebrain, midbrain, hindbrain and dorsal spinal cord) up to at least 36hpf (Girard et al., 2001). One clone isolated from the 5dpf stage showed very low sequence similarity (18bp outside the primer sequences) to *sox31*. With an Expect value (E) of 9.1 it is most likely a cloning artefact. It means that by chance, one might expect to see 9.1 sequences from the database to show a similar extent of identity.

From the 24hpf stage, a *sox2*-like homologue was isolated, which has not been described yet in zebrafish. It belongs to subgroup B1, like *sox1*, *sox3* and *sox19*. In chicken, *Sox2* is expressed primarily in the CNS, but also in spinal, sympathetic and cephalic ganglia, Schwann cells, putative parasympathetic nerves, branchial arches, nasal epithelium and both neural retina and lens (Uwanogho et al., 1995; Uchikawa et al., 1999).

An HMG box with high nucleotide identity to *Sox4* homologues (subgroup C) was cloned once from the 5dpf stage and twice from the adult stage. Again, a zebrafish *sox4* has not been published yet and the time and place of function is unknown. Three likely members of subfamily D were identified, which showed similarities to *sox5* (19hpf stage), *sox6* (19hpf embryos and 5dpf larvae) and *sox13* (19hpf and adult tissues). All three genes represent novel zebrafish homologues. In mice, *Sox5* and *Sox6* are involved in spermatogenesis and chondrogenesis (Lefebvre et al., 1998). *Sox13* is expressed in mouse embryos in developing arterial walls, inner ear and thymus and also in the thymus of adult mice (Roose et al., 1998). Consistent with this, we isolated a *sox13*-like homologue from 19hpf stage zebrafish embryos and adult tissue.

An HMG box with similarities only to a sea urchin *SoxB2* (Genbank accession number AF157388) was isolated from the 19hpf stage. The region of identity was small (23bp outside primer sequences) and no similarities to other *Sox* genes were found. *SoxB2* clusters with members of subgroup B2 and thus might represent either a *Sox14* or *Sox21*

orthologue (Bowles et al., 2000). However, with an Expect value E of 1.3, this fragment might only represent a cloning artefact.

Finally, we identified the *sox10*-like HMG box only once, from cDNA prepared from 19hpf zebrafish embryos. As described earlier, at the equivalent stage in mouse, *Sox10* is expressed widely in premigratory and migrating neural crest cells and the otic vesicle (Herbarth et al., 1998; Southard-Smith et al., 1998). Thus, it was not surprising to isolate it from tissue of this stage.

In conclusion, the HMG box clones obtained by RT-PCR generally fitted the available expression data of those *Sox* genes in the stages investigated.

Although unnecessary for the progress of this study at the time, it would indeed be interesting to continue this experiment with the aim of looking for a second *sox10* homologue, derived from the fairly recent genome duplication in ray-finned fish (Amores et al., 1998; Postlethwait et al., 1998). Any duplicated gene would have had to diverge in order to be maintained in the genome. Under the classical model for the evolution of duplicated genes, a duplicate is only preserved if a mutation creates a new function. Since this event is likely to be much rarer than any deleterious mutation, this model could not explain the high number of duplicates observed. An alternative model, called the Duplication-Degeneration-Complementation model (DDC), may explain the frequent maintenance of both duplicates (Force et al., 1999). It hypothesises that a member of the duplicated pair may undergo Nonfunctionalisation, in which it acquires a null mutation; Neofunctionalisation, in which a mutation gives rise to a new gene function; or Subfunctionalisation, in which the duplicates accumulate degenerative mutations in different regulatory elements. By complementing each others function, both duplicates are maintained. This in turn might even increase the chances of acquiring a novel gene function in either of them. Indeed, what is often observed is that the original role of the ancestral gene is temporally or spatially shared by the two paralogues. The zebrafish duplicates *engrailed1* (*eng1*) and *engrailed1b* (*eng1b*) are presented as one of the examples illustrating this subfunctionalisation. *Engrailed1* (*EN1*) in mouse and chicken is expressed in the pectoral appendage bud and in certain neurons of the hindbrain and spinal cord (Joyner and Martin, 1987 cited in Force et al., 1999). In zebrafish these two regions of expression are split between *eng1* and *eng1b*, with *eng1* expressed in the pectoral appendage bud and *eng1b* in neurons of the hindbrain (Force et al., 1999). Mammalian *Sox9* is expressed in testes and is involved in sex determination (Foster et al., 1994; Wagner et al., 1994; Kent, 1996). In zebrafish,

*sox9a* is expressed in testes, whereas *sox9b* is found in ovaries (Chiang et al., 2001). This change of expression pattern could indicate a new role adopted by the second paralogue.
It is impossible to predict, which subfunction a second *sox10* paralogue might have developed during evolution.

It remains untested whether the alternative cloning strategy involving the design of *sox10*-specific primers might have been a better method to use. Potentially, this specificity for *sox10* could have lead to the successful cloning of a *sox10*-like HMG box in even shorter time. However, the design of such primers might have been difficult since only sequences of mammalian *Sox10* homologues were available at the time. With the addition of a recently published chicken *Sox10* and a likely fish *sox10* homologue, the design of such degenerate *sox10* specific primers in highly conserved regions would be the cloning strategy of choice, if the procedure were to be performed now.

## Isolation of a full length *sox10* homologue

Once the HMG box had been cloned, extension of the sequence with the SMART RACE PCR Amplification Kit (Clontech) proved straightforward.
A 3'RACE product, clone B2 or C2, was obtained at the first attempt, although primer N3GSP had misprimed within the *sox10* sequence approximately 1000bp further downstream. It might be that the PCR conditions were not appropriate to amplify a product spanning the entire region 3' to the HMG box (about 2.8kb). In any case, shorter products are usually favoured. Alternatively, the cDNA population might have contained a lot of truncated reverse transcripts. With the specific annealing site missing, the primer might have been forced to bind to a location with lower sequence homology further downstream. Thus, an additional round of gene specific amplification was required to complete the 3' sequence. This overlapping fragment, represented by clones 15 and 20, linked the HMG box clone SoxS4 to the 3'RACE product, clone B2.
Clone 15 differed from clone 20 only by an insertion of 109bp within the 3'UTR. This insertion might represent a splicing variant or less likely a PCR artefact. A PAC clone containing the equivalent genomic region was sequenced and confirmed the presence of this 109bp insert in the *sox10* gene (T. Carney, unpubl. data). Thus, the insertion was included in the *sox10*-like cDNA.

Some perseverance was necessary to establish conditions to clone the 5'RACE fragment. Two clones, 32 and 62, were isolated, which were identical to each other apart from the N-terminal 26bp extension in clone 32. This might have been caused by mispriming of the SMART oligo on a series of cytosines further downstream in the *sox10* sequence.

Both the results from the BLASTX search and an unrooted maximum-likelihood tree consistently suggested that the *sox10*-like cDNA represented a true *sox10* homologue. As was demonstrated in Figure 3.17, the newly isolated member of the *Sox* gene family clearly clustered with other *Sox10*s instead of *Sox9* or *Sox8*s. In fact, an alignment with the two zebrafish *sox9* paralogues, both closely related to *sox10*, only exhibit an overall amino acid identity of 40 and 42.6%, respectively, whereas they show 59.6% identity to each other (Table 3.3)

The zebrafish Sox10-like predicted protein sequence also shows consistently higher identity to all published Sox10s (approximately 60%) than to Sox9s (approximately 43%) and even less homology to Sox8s (approximately 41%). Together these results strongly support the hypothesis that the cDNA indeed represents a true *sox10* homologue. In accordance with the zebrafish gene nomenclature it will be referred to as *sox10* (protein as Sox10) in subsequent chapters of this thesis.

We made no attempt to test whether the cloned *sox10* cDNA represented a full length clone. The SMART RACE PCR kit is designed to isolate full length 5' and 3'RACE fragments. Indeed, the presence of a 28 nucleotide poly-A tail strongly indicates a complete 3'end. However, to investigate the true 5' end of the mRNA transcript, a primer extension experiment is necessary, whereby the beginning of the 5'UTR is reverse transcribed with a labelled, gene specific primer. Once the genomic sequence (PAC clone) of that region and the length of the labelled product is known, the first nucleotide of the transcript can be determined. However, for the purposes of this study, a full length coding region was most important and thus, such primer extension studies were omitted.

The start of translation was determined by sequence alignment with other Sox10 proteins. The first 30 residues show significant conservation in all known Sox10 proteins and thus the corresponding initiator methionine in zebrafish was easily identified (Figure 3.16). Also, no alternative in frame methionine codon is present further upstream. Southard-Smith investigated an alternative methionine in mouse further upstream as a potential start of translation by aligning the mouse, human and rat

sequences (Southard-Smith et al., 1999). The region exhibited 90% nucleotide identity between mouse and human 5'UTR indicative of structural relevance. However, a frameshift present in the human sequence introduced a stop codon between the two possible initiator sites, thus eliminating the possibility of usage of alternative methionine further upstream.

The 485 residues representing the coding region of the *sox10* cDNA extend from nucleotides 376 to 1833 (Figure 3.15). The only significant difference from other Sox10 homologues C-terminal to the HMG domain is the insertion of 15 amino acids rich in Glycine, G (5 out of 15) and Serine, S (4 out of 15) (see Figure 3.16). The function of this region is unknown. The HMG domain (residues 106-185) is the most conserved region showing 95% amino acid identity to mammalian Sox10 homologues (Figure 3.16). It is preceeded by 38 residues that are highly conserved within members of the subgroup E. Zebrafish Sox10, for example, shows 83.8% amino acid identity to all known Sox10 homologues in this region. Their function still requires investigation.

An N-terminal domain, called the synergy domain, is poorly defined. Kuhlbrodt and colleagues (1998a) tested rat *Sox10* deletion constructs in *in vitro* luciferase reporter assays and concluded that both the HMG domain and the first 89 residues were required for synergistic interaction of Sox10 with other partner transcription factors such as Tst-1/Oct6/SCIP. The N-terminal domain of Sox10 shows a low degree of conservation compared to mammalian Sox10 proteins. Thus, residues 1-89 exhibit 42.5% amino acid identity between zebrafish and mouse, but 100% identity between human and mouse synergy domains (Figure 3.16).

The transactivation domain (TA domain) is also not well defined. Deletion constructs fusing various parts of the *SOX10* 3' end to the GAL4 DNA binding domain were tested for *in vitro* ability to transactivate a GAL4-dependent CAT-reporter plasmid (Pusch et al., 1998). Only the region between residues 377-466 produced a 29-fold CAT induction and thus exhibited transactivational abilities. This region is highly conserved in mammals (Figure 3.16) and also corresponded approximately to the transactivation domain previously identified in SOX9 (Suedbeck et al., 1996). The same region in zebrafish Sox10, residues 395-485, is not as highly conserved between zebrafish and human or chicken Sox10 (75% amino acid identity to either of them, Figure 3.16B). However, the high identity of 87.2% within the last 47 residues in all Sox10 homologues might point to a functionally important part of this domain.

A region rich in prolines (P, 5/13 residues) and threonines (T, 4/13 residues) and completely conserved across all subgroup E homologues is located between residues 240-252. Bondurand et al. (2000) have also mentioned such a domain in approximately the correct location and it would be interesting to investigate its function.

In conclusion, we were able to show the cloning of a true *sox10* homologue from zebrafish by RT-PCR and RACE-PCR and confirm its identity by phylogenetic analysis and sequence alignments. In the following chapter we will ask whether *sox10* is expressed in the neural crest, and whether expression in *cls* mutant embryos is disrupted, as might be expected if *cls* encodes *sox10*.

# Chapter 4 – Expression of *sox10* in wild-type and *cls* mutant embryos

## *Introduction*

In the previous chapter we have described the cloning of a *sox10*-like homologue. We could further show by sequence alignment and phylogenetic analysis that this *Sox* gene was a true *sox10* homologue in zebrafish.

We next asked whether this *sox10* homologue was expressed in the correct places to be a strong candidate for the *cls* gene. Cell autonomy studies of the *cls* phenotype indicated that *cls* would be expressed in the premigratory neural crest (Kelsh and Eisen, 2000b). Likewise, the strong ear phenotype in *cls* mutants would be consistent with expression in the otic vesicle. These aspects are observed in mammalian *Sox10* expression patterns. Furthermore, the expression pattern of *sox10* was expected to be disrupted in homozygous *cls* mutant embryos.

## *Sox10* expression during mouse, human and chicken development

The *Sox10* HMG box was first isolated from a mouse embryo cDNA library by degenerate PCR amplification (Wright et al., 1993). In the course of identifying the product of the *Dom* locus by a positional cloning approach, the full cDNA sequence in mouse was isolated by two independent research groups (Herbarth et al., 1998; Southard-Smith et al., 1998). Based on amino acid similarities within the conserved HMG domain and characteristics outside such as the presence of specific functional domains and conservation of intronic sites, *Sox10* was classified as a member of subgroup E together with *Sox9* and *Sox8* (Bowles et al., 2000). At the same time, rat *Sox10* was partially cloned by RT-PCR and this HMG box fragment used to isolate the full length clone from a Schwann cell cDNA library (Kuhlbrodt et al., 1998a). Phenotypic similarities between the $Sox10^{Dom/+}$ mouse and patients with Waardenburg Shah Syndrome (WS4) led to the proposal that WS4 might result from *SOX10* mutations. Consequently, *SOX10* was isolated from a human brain cDNA library using the rat *Sox10* cDNA as a probe (Pingault et al., 1998a). Only a year later, the cloning of a chicken *Sox10* homologue was reported (Schneider et al., 1999; Cheng et al., 2000).

### *Sox10 expression in the neural crest and its derivatives*

Although there are small species-specific differences in *Sox10* expression, the overall pattern appears largely conserved. Thus, in mammals and chicken, *Sox10* is strongly expressed in premigratory neural crest cells soon after emergence from the neural tube. In mouse and humans, the onset of *Sox10* (*SOX10*) expression has not been studied in detail. *Sox10* transcripts were first detected in E8.5 mice (5-12 somites) in a region corresponding to the edges of the neural plate, adjacent to areas which have just closed to form a neural tube. These cells most likely represent neural crest cells emerging from the neural tube (Kuhlbrodt et al., 1998a). In expression studies in chicken, *Sox10* transcripts were first detected just after *Slug* at HH (Hamburger & Hamilton) stage 7/8 (3 somites) in the folding neural plate, but not in the lateral region from which the neural crest delaminates. However, consistent with the mouse data, strong expression was observed in cranial and trunk neural crest slightly later, from HH stage 9 (~6-7 somites; Cheng, Cheung et al. 2000). The earliest stage investigated in human embryos was 4 weeks, the equivalent to mouse E10. *SOX10* expression was seen in the premigratory crest area (Bondurand et al., 1998a).

Furthermore, *Sox10* appears to be expressed in at least some migrating neural crest cells on either pathway and is maintained predominantly in glial-type cells of the PNS and CNS, whereas *Sox10* is down regulated in differentiating neuronal derivatives. In mouse, *Sox10* expressing neural crest cells were seen to migrate on both the medial pathway at E9.5 (Pusch et al., 1998) and on the lateral pathway until at least E12.5. The latter cells correspond to pigment cell precursors (Southard-Smith et al., 1998). Some migrating *SOX10*-positive neural crest cells were observed in 4 week human embryos (mouse E10) in the region of forming glosso-pharyngeal ganglia.

In contrast to mouse and chicken, *SOX10* was also detected in some derivatives of the cephalic neural crest in 6 week human embryos (mouse E14), which later form cartilaginous rudiments of several nasal bones (Bondurand et al., 1998a). Maintenance of *Sox10* expression in such ectomesenchymal fates was not observed in mouse (Kuhlbrodt et al., 1998a).

*Sox10* expression in the mouse peripheral nervous system was seen at E9.5 in DRG, trigeminal and the cochlear-vestibular ganglion (Kuhlbrodt et al., 1998a). By E10.5, *Sox10* staining was associated with all cranial ganglia, cranial and spinal nerves (motor and sensory nerves) and sympathetic ganglia (Pusch et al., 1998; Southard-Smith et al.,

1998). Furthermore, from E12.5 enteric ganglia in the trunk exhibited hybridisation signals (Kuhlbrodt et al., 1998a; Southard-Smith et al., 1998). All these crest derivatives were also strongly labelled by *SOX10* in human embryos (Bondurand et al., 1998a). In chicken, it was shown that *Sox10* expression was not maintained in neural crest cells undergoing neuronal differentiation as judged by expression of the Hu epitope, for example in the centre of peripheral ganglia in head and trunk (Cheng et al., 2000). This is consistent with findings of non-overlapping *Sox10* signals and the neuronal marker SorLA in the developing trigeminal ganglion in mouse (Kuhlbrodt et al., 1998a). Since high expression levels of *Sox10* were found in cultured Schwann cells, these *Sox10*-positive cells surrounding peripheral ganglia and along cranial and peripheral nerves most likely corresponded to satellite glia cells and Schwann cells, respectively (Kuhlbrodt et al., 1998a). It was suggested that some of these *Sox10*-positive cells in DRG might also represent undifferentiated neural crest cells since they did not express glial markers and as differentiated neurons would have turned off *Sox10* (Britsch et al., 2001).

### *Sox10 expression outside the neural crest*

Aside from the neural crest, it was suggested that *Sox10* in chicken was expressed in neuronal and glial precursors during early CNS development. Thus, the entire ventricular zone was labelled at HH26 (~5 days), which contains mostly neuronal precursors. At later stages expression was absent in Hu-expressing neuronal cells and became restricted to oligodendrocytes (Cheng et al., 2000). In mice, labelling intensities in the CNS increased from E12.5 to E18.5 drastically. Expression was observed throughout the brain corresponding to forming oligodendrocytes, but never in brain nuclei (Kuhlbrodt et al., 1998a). This was interesting since in contrast to mice, *SOX10* was also observed in major brain nuclei and in the cerebral cortex in both fetal and adult human CNS (Bondurand et al., 1998a).

High levels of *Sox10* are noted in the otic vesicle from E9.5 in mouse (Pusch et al., 1998; Southard-Smith et al., 1998) and between the 5$^{th}$ (corresponding to ~E11 in mouse) and 6$^{th}$ week (~E14) during human development (Bondurand et al., 1998a). The otic placode in chicken is labelled very early from HH stage 9 (~6-7 somites) and expression there is maintained strongly at least up to HH stage 27 (~5.5 days) in non-neuronal cells as shown in double labelling with anti-Hu antibody.

*Sox10* transcripts were also observed in the developing pineal gland of chicken at HH stage 26 (~5 days; Cheng et al., 2000).

## *Sox10* mutant mice have altered *Sox10* expression patterns

The *Dom* mouse was the first *Sox10* mutant to be identified. An insertion of a single guanine residue at nucleotide 929 causes a frameshift, which leads to the generation of 99 novel amino acids before premature truncation of the protein (Southard-Smith et al., 1998). Its phenotype has been described in the Introduction chapter.

### *Sox10 expression patterns in Dom mutants in the neural crest*

*In situ* hybridisation with *Sox10* in $Sox10^{Dom/Dom}$ mice revealed that expression was present in premigratory neural crest cells at E9.5 in a wild-type pattern. Expression persisted in migrating neural crest cells in caudal regions until E10.5 although these neural crest cells were delayed in their migration. By E10.5, cranial ganglia and cells associated with their projections showed strongly reduced *Sox10* levels. *Sox10*-positive cells in the gut were absent and were reduced in the sympathetic primordium (Britsch et al., 2001). Melanoblasts, normally expressing *Sox10* and *Dct*, an early melanoblast marker, were absent. At E11.5, *Sox10* transcripts were only detectable in the caudal-most neural crest cells, which, in developmental terms, are the youngest. At the same time, cells in DRG and in a location consistent with migrating neural crest were undergoing significant apoptosis as seen by TUNEL staining (Terminal deoxynucleotidyl transferase-mediated dUTP nick end-labelling) (Southard-Smith et al., 1998). In contrast to wild-type embryos, *in situ* hybridisation on sections of E14.5 $Sox10^{Dom/Dom}$ mutant embryos demonstrated a lack of enteric precursors in their intestines and a reduction in $Sox10^{Dom/+}$ siblings (Herbarth et al., 1998).
These results demonstrate disrupted *Sox10* expression in $Sox10^{Dom/Dom}$ mutants and are consistent with the defects in neural crest cell lineages observed in the $Sox10^{Dom/+}$ mouse. Apoptosis of neural crest cells in hetero- and homozygous $Sox10^{Dom}$ mutants appears to cause a reduction of pigment cells and enteric ganglia (Southard-Smith et al., 1998).

### *Sox10 expression patterns in Sox10$^{Dom}$ mutants outside the neural crest*

In contrast to that in neural crest derivatives, expression in the otic vesicle, the CNS and cells associated with developing cartilage persisted in *Sox10$^{Dom/Dom}$* mutants (Britsch et al., 2001).

## *Results*

### The *sox10* expression pattern in wild-type zebrafish

We asked whether the expression pattern of zebrafish *sox10* was equivalent to the one reported for mammalian and chicken *Sox10*. To study the cloned gene's expression pattern in wild-type zebrafish, we performed mRNA *in situ* hybridisations on whole mount embryos of various developmental stages.

Thus, we synthesised an antisense mRNA *in situ* probe against the *sox10* sequence C-terminal to the HMG box (Materials and Methods). The 1913bp 3'RACE clone 20 was used as a template. It contained 988bp of coding sequence beginning with the 85bp 3'end of the HMG box domain and including 925bp of the 3' untranslated region (3'UTR). Prior to RNA synthesis with RNA polymerase T7, the plasmid containing clone 20 was linearised with *Sal* I, which also removed the 3'end of the HMG box sequence. This was considered important in order to prevent possible cross-hybridisation due to the high degree of sequence conservation within this domain between different *Sox* genes (Figure 4.1).

Zebrafish embryos were collected from heterozygous *cls$^{m618}$* crosses, staged according to (Kimmel et al., 1995) and fixed at appropriate stages for a timecourse. *In situ* hybridisations on batches of 20-30 embryos per staining reaction were carried out with 1:200 dilutions of the *sox10 in situ* probe (Materials and Methods).

### *sox10 expression in the neural crest and its derivatives*

We asked whether *sox10* expression was maternally transmitted. However, no expression was detected at the 4 cell stage, sphere stage (4hpf), shield stage (6hpf) and at 90% epiboly (9.5hpf).

First expression of *sox10* was observed in 2 somite stage embryos (approximately 10.5 hpf) at the lateral edges of the neural plate (Figure 4.2A). This region corresponded to the emerging neural crest. At 5 somites (12hpf) this expression domain had broadened

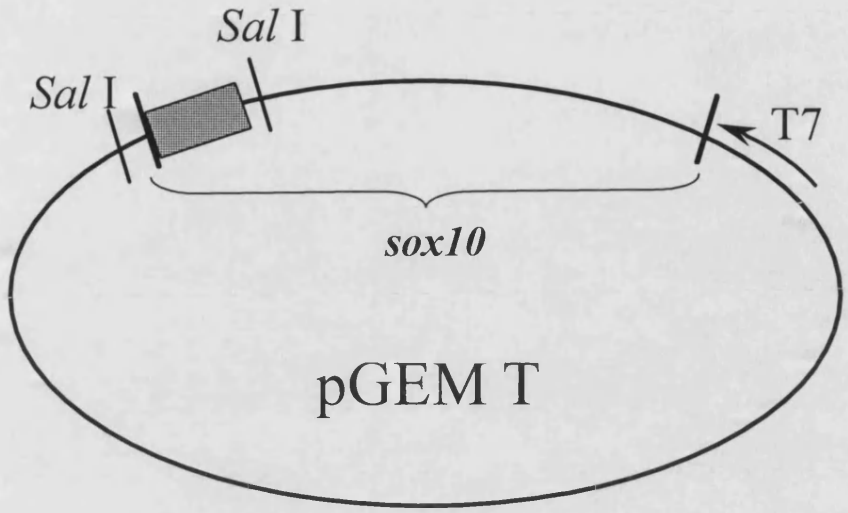
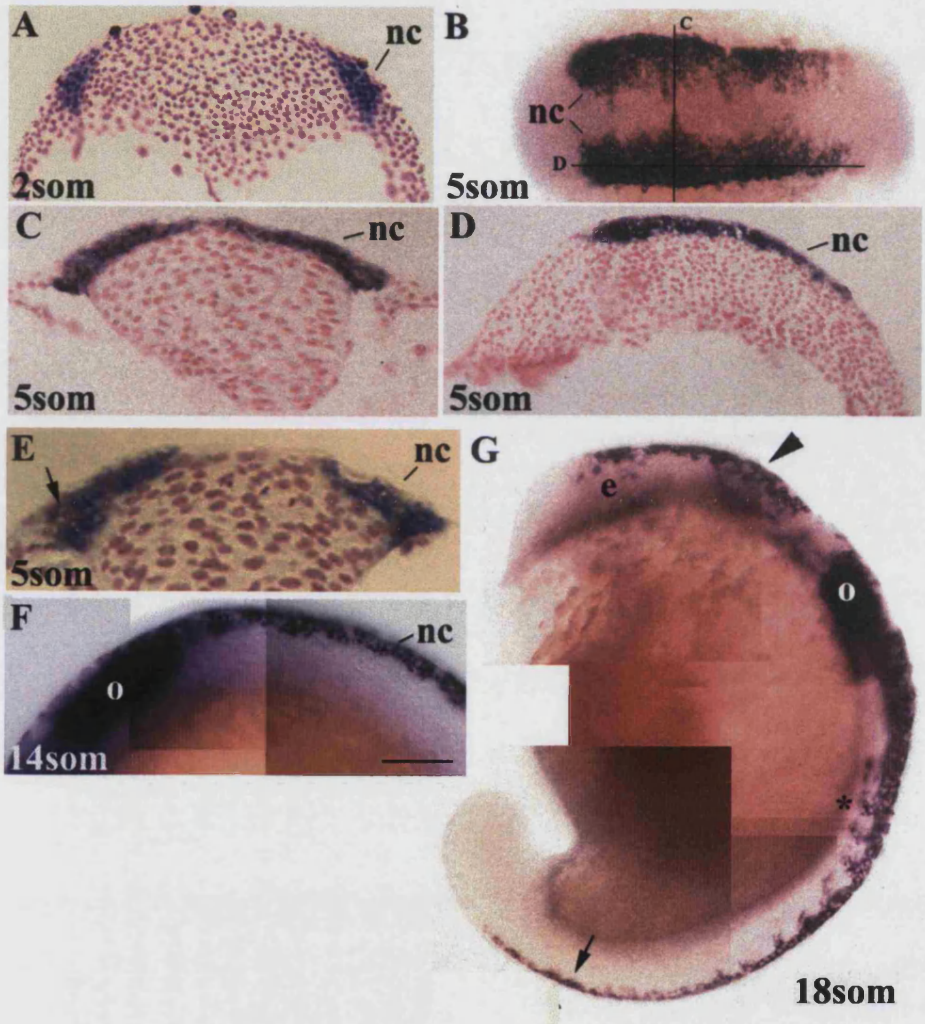



**Figure 4.1:** Preparation of the *sox10 in situ* hybridisation probe

Clone 20 was used as a template for antisense RNA transcription. The clone contains the *sox10* cDNA fragment 3' to the HMG box ( ▓ ), including 85bp of HMG box sequence and 925bp of 3'UTR. Prior to probe synthesis, the plasmid was linearised with *Sal* I at the 5' end of the cDNA, which also removed the HMG box sequence. This was thought to be necessary to prevent possible cross hybridisation to other HMG box containing genes. RNA polymerase T7 was then used to synthesise 1825bp antisense mRNA transcripts.

**Figure 4.2:** *sox10* expression in wild-type and *cls* embryos up to the 18 somite stage Expression patterns in *cls* mutant embryos were indistinguishable from wild-type siblings up to the 18 somite stage. (A) *sox10* expression was first observed in emerging neural crest cells in the lateral neural plate from the 1-2 somite stage. By the 5 somite stage, cranial crest cells had formed two broad stripes of cells (B). In the example of a transverse (C) and longitudinal (D) section shown (plane of section indicated in (B), all neural crest cells seemed to express *sox10*. However, very occasionally, a single *sox10* negative cell amongst labelled cells was observed (black arrow, E). Premigratory trunk neural crest cells expressed *sox10* in 14 somite stage embryos (F) and in 18 somite stage embryos (black arrow, G). Furthermore, strong expression was maintained in cranial neural crest cells (black arrow head, G) and migratory trunk neural crest cells (asterisk, G). The otic placode (o) was stained strongly from the 11 somite stage (data not shown), in 14 and 18 somite embryos (F, G). In transverse sections (A, C, E) dorsal is up and in the longitudinal section (D) dorsal is up and anterior to the left. Wholemount (B) is shown in a dorsal view with anterior to the left, whereas wholemounts (F, G) are lateral views with anterior to the left (F) and anterior up (G). e, eye; nc, neural crest; o, otic placode. Scale bar: 145μm in A-B, D; 60μm in C; 50μm in E; 95μm in F; 75μm in G.


A
nc
2som
B
C
nc
D
5som
C
nc
5som
D
nc
5som
E
nc
5som
F
o
nc
14som
G
e
o
*
18som

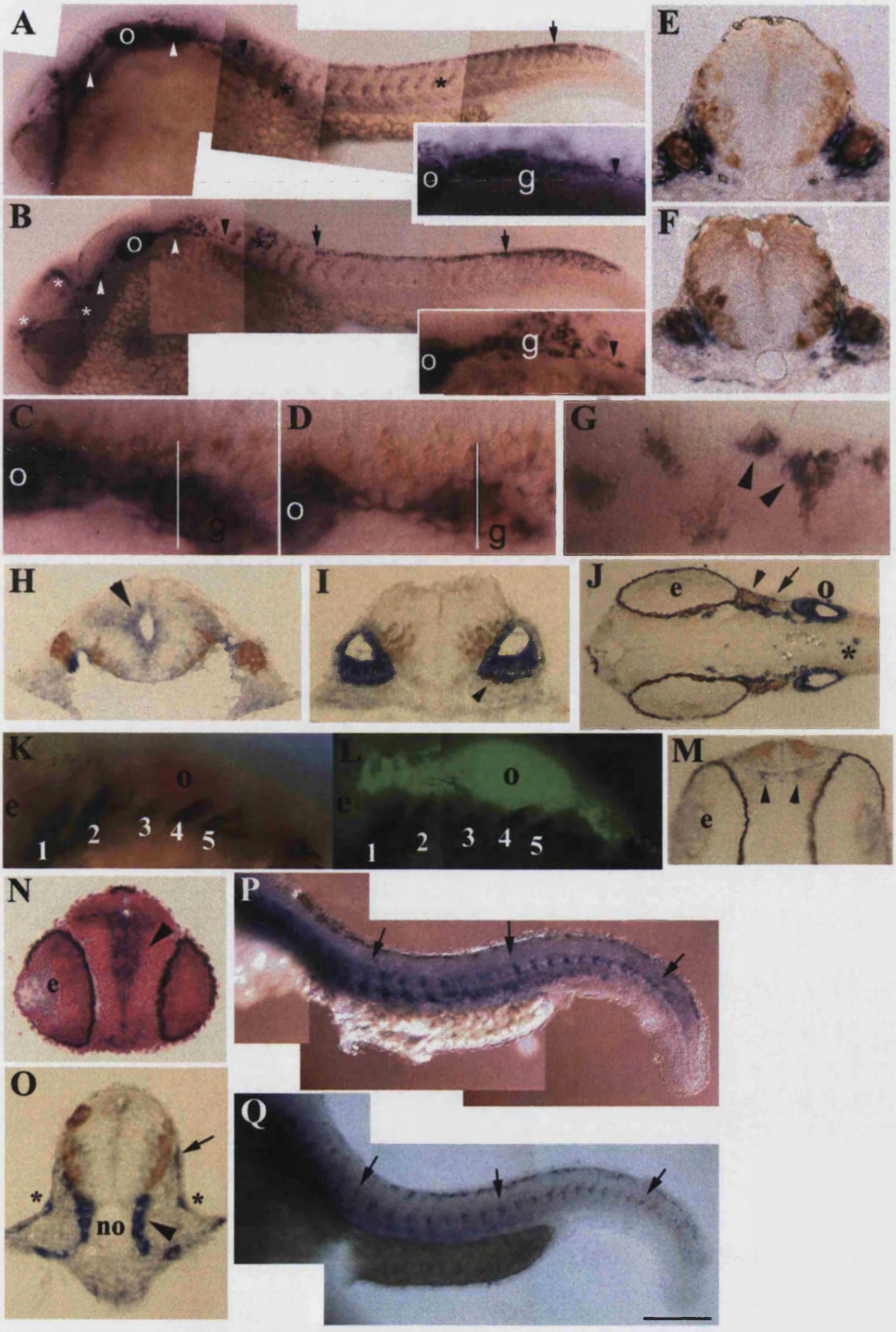
and neural crest cells appeared to have started to converge dorsally (Figure 4.2B, C). We asked whether all neural crest cells were *sox10*-positive and thus prepared sections of 2 somite and 5 somite embryos. Transverse and longitudinal sections through the neural keel showed that most, if not all, neural crest cells were labelled with *sox10* (Figure 4.2C, D).

The strong expression of *sox10* in premigratory neural crest cells was maintained and extended caudally during later stages of somitogenesis. At approximately 12-14 somite stage (15-16hpf), cranial neural crest cells just caudal to the eye start to migrate away from their premigratory positions (Schilling and Kimmel, 1994). Trunk neural crest cells start to enter the medial pathway at approximately the 14-16 somite stage (16-17hpf) and the lateral pathway approximately 4 hours after the first cells started to migrate medially at this axial level (Raible et al., 1992). Consistent with the literature, trunk neural crest cells remained in the migration staging area above the somites until at least the 14 somite stage (16hpf, Figure 4.2F). By the 18 somite stage (18hpf) more caudally located neural crest cells still remained premigratory, but at least some *sox10*-positive neural crest cells were seen to migrate on the medial pathway in the anterior trunk region (Figure 4.2G). At 24hpf, anterior trunk neural crest cells labelled with *sox10* were seen on the medial and lateral pathway (Figure 4.3O).

Melanophores were the first neural crest derivatives of the pigment cell lineage to differentiate. We asked whether *sox10* expression was maintained in differentiated melanophores. By 30hpf, migrating melanophores were visible on the medial and lateral pathway, but *sox10* expression in these cells was not readily detectable. Occasionally, a melanophore dorsal to the neural tube was seen to weakly express *sox10* (Figure 4.3G). Since only pigment cell precursors are known to migrate on the lateral pathway (Raible and Eisen, 1994), this finding indicated that *sox10* was turned off rapidly in differentiating pigment cells. Consistent with this, expression in pigmented xanthophores was never observed.

Derivatives of the peripheral nervous system were first seen around 24hpf. Cells expressing high levels of *sox10* accumulated in the region of forming cranial ganglia and extended caudally from the posterior lateral line ganglion by 24hpf (Figure 4.3A). By 36hpf, *sox10*-positive cells were observed to be clustered around the trigeminal, anterior and posterior lateral line ganglia (Figure 4.3J). We asked whether these *sox10*-positive cells were glial and/or neuronal derivatives. The expression patterns observed were reminiscent of *in situ* hybridisation labelling studies with *fkd6*, a glial marker

**Figure 4.3:** *sox10* expression in wild-type and *cls* mutant embryos at 24-36hpf stages
From 24hpf, some differences between the wild-type (WT) and *cls* expression patterns are observed.
(A) In 24hpf WT strong *sox10* expression is seen in cranial ganglia (white arrowheads) and along the posterior lateral line nerve (PLLn, A, black arrowhead, enlarged in inset; in transverse section in O), in premigratory (A, black arrow) and migrating crest cells (A, black asterisk) on both medial (O, black arrowhead) and lateral pathways (O, black arrow). Cells marked by an asterisk in O might also migrate laterally or be Schwann cells on the PLLn. Strong expression is seen in the otic vesicle (o) in the epithelial layer (I). Note the forming acoustic ganglion (black arrowhead, I). (B) In contrast, 24hpf *cls* embryos show much reduced labelling in the cranial ganglia (white arrowheads) and on the PLLn (black arrowhead, enlarged in inset). Crest cells in the head are clustered (white asterisk). The expression in the ear is comparable to WT. *sox10*-positive cells migrate normally in the rostral trunk (black asterisk), but some cells tend to cluster in premigratory positions (black arrows). (C, D) Double labelling with *sox10* (purple) and the pan-neuronal anti-Hu antibody (orange) shows strong expression associated with cranial ganglia (g) in WT, which is reduced in *cls* (D). Sections through the posterior lateral line ganglion (approximate position indicated by white line in C, D) in WT (E) and *cls* mutants (F) shows that *sox10* expression is restricted to satellite glia (purple) surrounding the neuronal (orange) centre of the ganglion. (G) *sox10* is down regulated rapidly in differentiating melanophores. At 31hpf some still show weak expression in the dorsal stripe (black arrowheads). (H) Weak *sox10* staining is seen in the ventricular layer of the hindbrain (black arrowhead) at 24hpf. (J, M, N) In 36hpf (J) and 40hpf (M, N) WT embryos, expression is maintained in satellite glia of all cranial ganglia (trigeminal, black arrowhead; anterior lateral line ganglion, black arrow), the otic vesicle (o) and in the ventricular layer in the brain (J, asterisk; M and N, black arrowheads). (K, L) Double *in situ* hybridisation with *dlx2* (K, purple) and *sox10* (L, green) in 29hpf stage embryos shows absence of *sox10* expression in developing branchial arches (1-5). (P, Q) In 35hpf WT and *cls* embryos, segmentally arranged clumps of 3-4 *sox10* positive cells (black arrows) are observed along the entire trunk and tail, adjacent to the dorsal notochord, presumably glia of dorsal root ganglia or Schwann cells on segmental nerves. The *sox10* expression pattern in this region of *cls* embryos is very similar to wild-types (Q).
Anterior is to the left in all lateral views (A-D, K, L, P, Q) and in dorsal views (G, J). In the dorsal view (M), anterior is up and in all transverse sections (E, F, H, I, N, O), dorsal is up. e, eye; g, posterior lateral line ganglion; no, notochord; o, otic vesicle. Scale bar: 165μm in A-B, P-Q; 50μm in C-D; 80μm in E-F, M-N; 90μm in G; 60μm in H-I, O; 110μm in J; 70μm in K-L.


A
o
*
o
g
B
o
*
o
g
E
F
C
o
g
D
o
g
G
H
I
J
e
o
*
K
e
o
1
2
3
4
5
L
e
o
1
2
3
4
5
M
e
N
e
P
O
*
*
no
Q

(Kelsh et al., 2000a). Thus, *sox10* expressing cells were thought to be developing satellite glial cells surrounding all cranial ganglia and Schwann cells associated with the posterior lateral line nerve. In accordance with this, *sox10*-positive cells in ganglia were not double-labelled with the pan-neuronal marker anti-Hu antibody, thus confirming their glial, not neuronal, identity (Figure 4.3C, E, H, J). However, without further tests we cannot exclude the possibility that some of these *sox10*-positive cells might also represent undifferentiated neural crest cells as suggested in mouse (Britsch et al., 2001). Furthermore, we could show in a double *in situ* hybridisation that *sox10*-positive cells on the lateral line nerve also expressed *fkd6* at 24hpf confirming their glial nature (data not shown). The expression in Schwann cells was maintained up to 60hpf (Figure 4.4G).

At 35hpf *sox10* expression was seen in segmentally arranged cells adjacent to the notochord (Figure 4.3P). These could represent undifferentiated neural crest cells migrating on the medial pathway or differentiated neurons or glia associated with DRG and segmental nerves. Since they were all found at the exact same dorsal ventral position, unusual for migrating cells, and appeared as cell clumps of usually 3-4 individuals, the latter was considered more likely. Double labelling with anti-Hu antibody or a glia marker such as *fkd6* will be required to investigate the specification status of these cells. *sox10* expression was absent from cells at this location shortly thereafter (approximately 36-40hpf; data not shown).

Enteric nervous system precursors expressed *sox10* as early as 24hpf (R. Kelsh, unpubl. observation). By 60hpf, they were strongly labelled and surrounded the gut epithelium (Figure 4.4J).

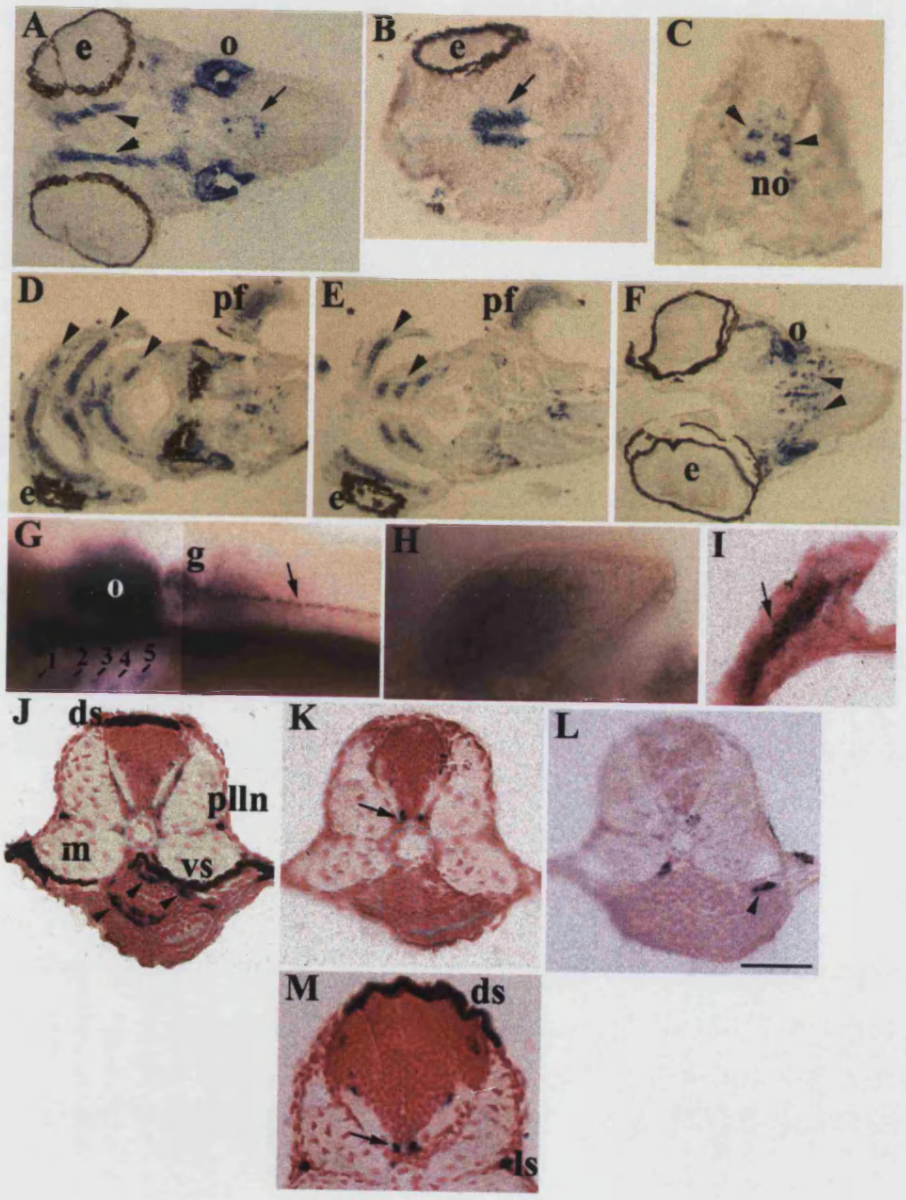
We asked whether any ectomesenchymal neural crest derivatives (medial fin mesenchyme and craniofacial cartilage) expressed *sox10*, even though *cls* embryos do not show any overt defects in these structures. We failed to detect any *sox10* expression in fin mesenchyme and craniofacial cartilage precursors. We used double *in situ* hybridisation to label craniofacial cartilage precursors with *dlx2*, but the regions of *sox10* expression did not overlap with *dlx2* at 29hpf (Figure 4.3K). However, at 48hpf and 60hpf, strong *sox10* staining was observed in the differentiating jaw cartilage (Figure 4.4A, D, G).

**Figure 4.4:** *sox10* expression in 48-60hpf wild-type and *cls* embryos

*sox10* expression in *cls* embryos in differentiating craniofacial cartilage, pectoral fins and CNS is indistinguishable from wild-types (WT).

(A-C) At 48hpf, *cls* embryos show expression in the forming jaw cartilage (A, black arrowheads), the otic epithelium (A, o) and in the ventricular layer (A,B, black arrow). (C) Strong expression is also seen in oligodendrocytes in a transverse section of the spinal cord of a *cls* embryo (black arrowheads). Staining persists until 60hpf in *cls* embryos in the jaw cartilage (D, E, black arrowheads), in oligodendrocytes (F, black arrowheads) and in cartilage precursors in the pectoral fins (pf, D, E, H). (I) A transverse section through a 60hpf *cls* pectoral fin demonstrates the typical "stack of coins" structure of cartilage precursor cells (black arrow). (F, G) Expression in the otic vesicle (o) at 60hpf is strong in wild-types (G). In contrast, *sox10* expression in the otic vesicle (o) in *cls* mutants is reduced by 60hpf (F). (G) Schwann cells on the posterior lateral line nerve (PLLn, black arrow) extending from the posterior lateral line ganglion (g) have maintained *sox10* in 60hpf WT embryos and expression is seen in branchial arches (1-5). (J-L) Melanophores (brown) have formed the dorsal (ds), ventral (vs) stripe and Schwann cells are located along the PLLn in WT (J), whereas both crest derivatives are absent in *cls* siblings (K, L). Likewise, numerous enteric precursors surrounding the wild-type gut lumen (J, black arrowheads) are usually completely absent in *cls* (K), but occasionally a single *sox10* positive cells is observed in this region (L, black arrowhead). (K, M) In contrast, oligodendrocytes (black arrows) in 60hpf WT (M) is unaffected in *cls* (K).

In longitudinal sections (A-B, D-F) and lateral views (G-H) anterior is to the left. In transverse sections (C, I-M) dorsal is up. ds, dorsal stripe; e, eye; g, posterior lateral line ganglion; ls, lateral stripe; m, muscle; no, notochord; o, otic vesicle; pf, pectoral fin; vs, ventral stripe. Scale bar: 165μm in A-B, D-F; 80μm in C; 95μm in G; 60μm in H-I; 65μm in J-M.


A
e
o
B
e
C
no
D
pf
e
E
pf
e
F
o
e
G
o
1 2 3 4 5
g
H
I
J
ds
plln
m
vs
K
L
M
ds
ls

### *sox10 expression outside the neural crest*

Aside from the neural crest, strong expression was observed in the otic placode and developing otic vesicle from the 11 somite stage onwards (Figure 4.2F,G, Figure 4.3A, C, I-K, Figure 4.4A, G). The cells of the otic epithelium were strongly labelled at all stages investigated (Figure 4.3I, J, Figure 4.4A).

Faint staining could be detected in the ventricular layer in the hindbrain from 24hpf (Figure 4.3H) and was seen additionally in the fore-and midbrain and the spinal cord at 36hpf (Figure 4.3J, M, N, Figure 4.4C). At 48hpf, *sox10*-positive cells were found in what appeared to be the subventricular layer in the forebrain (Figure 4.4B) and also started to disperse throughout the brain (Figure 4.4F). Consistent with *Sox10* expression in rodents (Kuhlbrodt et al., 1998a), this is likely to correspond to expression in developing oligodendrocytes. The staining in the floorplate was maintained at least up to 60hpf (Figure 4.4C, K, M).

At 24hpf weak *sox10* expression was observed in the developing pineal organ in a double *in situ* hybridisation with *fkd6* (Figure 4.6J, K). Interestingly, some cells appeared to express only *fkd6*, but their identity is unknown.

Finally, *sox10* expression was detected in the pectoral fins at 48hpf and 60hpf (Figure 4.4D, H, I). The staining was restricted to cells in the central part of the fin in a "stack of coins" arrangement (Figure 4.4I) typical for cartilage cells.

## Does *sox10* label all neural crest cells?

Analysis of the *colourless* phenotype showed that in contrast to non-ectomesenchymal derivatives, ectomesenchymal neural crest fates such as craniofacial cartilage and fin mesenchyme were unaffected (Kelsh and Eisen, 2000b; Dutton et al., 2001b). This prompted us to ask whether *sox10* was expressed in all premigratory neural crest cells or just in a subset, which might be specified to give rise to non-ectomesenchymal fates.

In an attempt to answer this question, we compared the expression pattern of *sox10* with *crestin* in a double RNA *in situ* hybridisation. *crestin* has been published as a pan-neural crest marker that labels all premigratory neural crest from approximately the 3 somite stage. It is then maintained in migrating neural crest cells until overt differentiation (Luo et al., 2001).

For this experiment, 6 and 12 somite stage embryos were chosen. By the 6 somite stage, premigratory cranial neural crest cells are known to be arranged in 6 tiers. The medial-

lateral position predicts the type of derivative which a neural crest cell will likely form (Schilling and Kimmel, 1994). In addition, it was demonstrated that 88% of all labelled neural crest cells were fate-restricted and gave rise to only a single cell-type (Schilling and Kimmel, 1994). We predicted differential gene expression depending on the fate-restriction of each crest precursor. Two research groups have investigated fate maps of cranial neural crest cell precursors, with slightly differing results. Schilling and colleagues found that ectomesenchymal derivatives arose from more medial tiers (3-6) together with some non-ectomesenchymal derivatives, but cells from tier (6) were of an exclusively cartilage and connective tissue nature (Schilling and Kimmel, 1994). In contrast, Dorsky and colleagues reported to have observed cartilage precursors from all tiers, even in the most lateral tier (1), which never gave rise to this cell-type according to Schilling (Dorsky et al., 1998). Thus, if *sox10* was not expressed in cartilage precursors, such *sox10*-negative cells are predicted to be seen mostly medially, but perhaps scattered across the entire neural crest region.

Embryos from a heterozygous $cls^{m618}$ cross were used in these double *in situ* hybridisations. The rationale behind this decision was to test whether the transcription of *crestin* was dependent on Sox10. If this was the case, one would expect to see differences in *crestin* expression in some embryos, presumably *cls* mutants. Unfortunately, we were unable to reproduce the published expression pattern of *crestin*. In our hands, *crestin* appeared to label all trunk neural crest cells in 6 and 12 somite stage embryos as published. However, expression of *crestin* in cranial crest cells appeared to be absent. Only by the 16 somite stage were a few scattered *crestin*-positive cells observed anterior to the otic vesicle and around the eye (Figure 4.5F, G). In comparison, a much greater number of *sox10*-positive cells were observed in the cranial region at this stage (data not shown). Intriguingly, a small cluster of cells in the centre of the otic vesicle was labelled with *crestin* (Figure 4.5G).

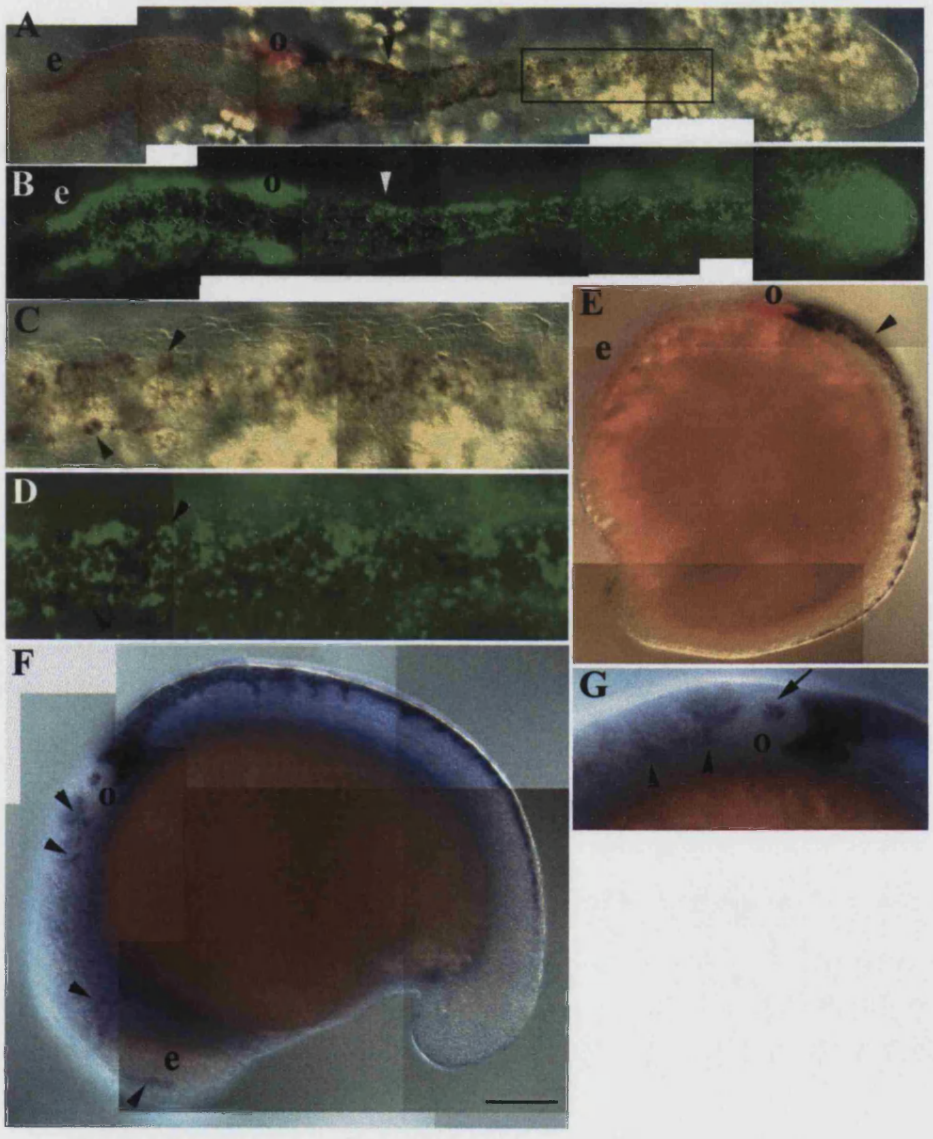
Thus, in 12 somite stage embryos, *sox10* and *crestin* staining broadly overlapped in the trunk posterior to the otic vesicle, but premigratory and migratory cranial neural crest cells almost exclusively expressed *sox10* (Figure 4.5A, B, E). We asked whether all neural crest cells in the trunk were labelled with both markers. We could not readily detect any $crestin^{+}/sox10^{-}$ cells in the premigratory position of the anterior trunk. This assessment was more difficult caudally, since premigratory cells in this region only weakly expressed either marker. Embryos were sectioned to try to confirm this result. Unfortunately, these sections were uninformative due to persistent technical problems in

**Figure 4.5:** Does *sox10* label all crest cells?

(A-E) Double RNA *in situ* hybridisation with *sox10* (brightfield, red, A, C, E; fluorescence, green, B, D, G) and *crestin* (purple) reveals incomplete overlap of expression patterns.

(A, B, E) In a 12 somite stage embryo the two markers show complete overlap in trunk neural crest cells (arrowhead) posterior to the otic placode (o), whereas anteriorly, cranial crest cells and the otic placode only express *sox10* (A, B, E, red). (C, D) *Crestin* and *sox10* transcription is initiated at approximately the same time. Even cells in further caudal regions (black arrowheads) in the enlarged region indicated by a black box in (A) express both genes. (F, G) *crestin in situ* hybridisation on 16 somite stage embryos shows some scattered *crestin* positive cells (black arrowheads in F and region enlarged in G) anterior to the otic placode (o) and around the eye (e), presumably cranial crest cells. A distinct patch of cells in the centre of the otic placode is labelled (G, black arrow).

In dorsal (A-D) and lateral views (F-G) anterior is to the left. In the lateral view (E) anterior is up. Scale bar: 260μm in A-B, E; 100μm in C-D; 75μm in F; 50μm in G.


A
e
o
B
e
o
C
D
E
o
e
F
o
e
G
o

the sectioning process. Using the standard sectioning method, the fluorescent *sox10* signal was lost.

*Crestin* and *sox10* expression in the neural crest is switched on at approximately the same time, at the 3 somite and 2 somite stage respectively (Dutton et al., 2001b; Luo et al., 2001). Consistent with the published data, even the developmentally youngest neural crest cells in the caudal-most region were labelled with both markers (Figure 4.5C, D). No obvious difference in *crestin* expression could be detected in any embryos. Thus, transcription of *crestin* did not seem to require regulation by Sox10.

The question whether *sox10* labelled all neural crest cells could not be answered with *crestin* for the reasons described above. Hence, we decided to compare the *sox10* expression pattern with that of *forkhead 6* (*fkd6*), which is widely expressed in premigratory neural crest from about 90% epiboly (Odenthal and Nuesslein-Volhard, 1998). Double *in situ* hybridisation was carried out on 6 and 14 somite stage embryos from a heterozygous $cls^{m618}$ cross. The 6 somite stage and use of *cls* mutant embryos was chosen for the same reasons outlined above. 14 somite stage embryos were available and thus used instead of the 12 somite stage for trivial reasons.

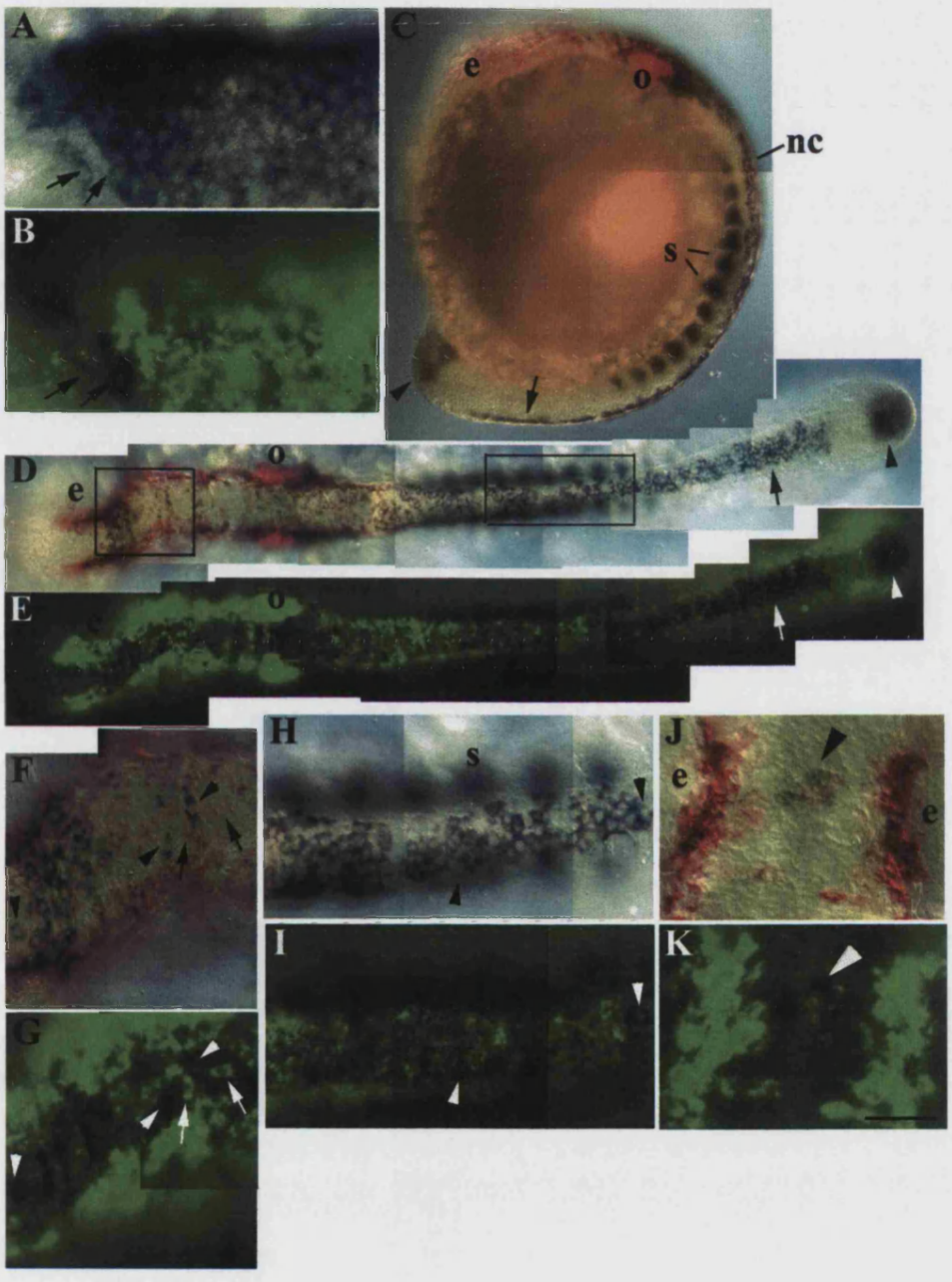
Labelling with *sox10* and *fkd6* revealed an extensive, but incomplete overlap of signals. Occasionally, a few $fkd6^{+}/sox10^{-}$ cells at the anteriormost edge of neural crest expression were observed (Figure 4.6A, B). Whether such cells were found elsewhere was impossible to assess in our *in situ* hybridisation studies. When several layers of neural crest cells were present, the coloured NBT precipitate (*fkd6* label) quenched the fluorescent signal (*sox10* label). Thus, the staining of each individual cell could not be determined. Sections of these embryos were also uninformative since the fluorescent signal was lost during the sectioning process.

At the 14 somite stage, a larger number of neural crest cells were observed that either expressed *sox10* or *fkd6*. The majority of these cells were found within the cranial neural crest cells (Figure 4.6D-G). *fkd6* expression is extinguished in migratory neural crest cells and by the 14 somite stage staining is also observed in somites and the tailbud (Figure 4.6C-D; Odenthal and Nuesslein-Volhard, 1998; Kelsh et al., 2000a). At this stage, most cranial neural crest cells have started their migration and thus it could be argued that $sox10^{+}/fkd6^{-}$ cells were migratory and had switched off *fkd6*. However, there was also a number of $fkd6^{+}/sox10^{-}$ neural crest cells arranged in rows above the brain especially posterior to the developing eye (Figure 4.6F, G). The degree of fate restriction of these *sox10*-negative cells is unknown, but they could represent

**Figure 4.6:** Does *fkd6* and *sox10* expression overlap completely in premigratory neural crest?

(A, B) Double *in situ* labelling of 6 somite stage embryos with *fkd6* (brightfield, purple, A) and *sox10* (fluorescent, green, B) shows incomplete overlap. A few cells (black arrows) on the anterior edge are exclusively labelled with *fkd6*. (C-I) 14 somite stage embryos double *in situ* labelled with *fkd6* (brightfield, purple, C-D, F, H) and *sox10* (fluorescent, green, E, G, I) are shown. (C) *fkd6* is never expressed in the otic placode (o) and is switched off in most cranial crest cells (red) as they start migration, whereas premigratory trunk crest cells posterior to the otic placode still express both markers. (C-E) *fkd6* also labels somites (s) and the tailbud (black arrowhead). The caudalmost neural crest cells (arrows) are *sox10*$^{-}$/*fkd6*$^{+}$ due to earlier onset of *fkd6* compared to *sox10*. (D, E) Scattered cells in the head, mostly posterior to the eyes (e), show differential expression of *fkd6* and *sox10*. The cranial region outlined (D, black box; enlarged in F, G) shows *fkd6*$^{+}$/*sox10*$^{-}$ cells (arrowheads) and *sox10*$^{+}$/*fkd6*$^{-}$ cells (arrows). Similarly, *sox10*$^{-}$/*fkd6*$^{+}$ cells (arrowheads) are seen in the posterior trunk region (D, black box above somites; enlarged in H, I). (J, K) Incomplete overlap between *sox10* and *fkd6* is also seen in the developing pineal organ (arrowhead) at 24hpf.

In dorsal views anterior is to the left (A-B, D-I) or anterior is up (J-K). In the lateral view (C) anterior is up. e, eye; nc, neural crest; o, otic placode; s, somite. Scale bar: 50μm in A-B, F-G; 135μm in C-E; 55μm in H-I; 30μm in J-K.


A
B
C
e
o
nc
s
D
e
o
E
o
F
G
H
s
I
J
e
e
K

premigratory cartilage precursors that will contribute to craniofacial structures not affected in *cls* mutants. The most newly formed neural crest cells were only labelled with *fkd6* as shown by the caudalmost regions of these embryos. This was consistent with the onset of *fkd6* transcription at 90% epiboly preceeding *sox10* at the 2 somite stage in the neural crest (Odenthal and Nuesslein-Volhard, 1998; Dutton et al., 2001b; Figure 4.6D, E). Occasionally, a *fkd6*$^{+}$/*sox10*$^{-}$ neural crest cell was observed in the trunk (Figure 4.6H, I). As in the head, it was difficult to distinguish between a true *sox10*$^{-}$ cell and such cells that had not yet turned on *sox10*. The majority of neural crest cells however appeared to be double-labelled.

Finally, *fkd6* expression in premigratory crest cells was not disrupted in *cls* mutants and thus *fkd6* did not appear to be regulated by *cls*. In *cls* mutants, fewer *fkd6* labelled Schwann cells were observed. However, this is likely to be due to a reduction in number of Schwann cells, rather than a reduction in *fkd6* expression levels.

## *sox10* expression is disrupted in *cls* mutant embryos

Pigment cells and enteric neurons and glia are almost absent in *cls* mutant embryos. Less severe defects are noted in other peripheral nervous system derivatives and the ear (Kelsh and Eisen, 2000b). Furthermore, *cls* was shown to act cell-autonomously at least in the pigment cell lineage (Kelsh and Eisen, 2000b). Thus, if *cls* encoded a *sox10* homologue, we expected to see a disrupted *sox10* expression pattern in tissues of *cls* mutants consistent with the *cls* mutant phenotype.

We performed RNA *in situ* hybridisation on embryos of a heterozygous *cls*$^{m618}$ cross on stages suitable for a timecourse. Siblings were carefully examined for consistent differences in expression patterns in a quarter of embryos corresponding to homozygous *cls* mutants.

### *sox10 expression patterns in cls mutants in the neural crest*

The pattern of *sox10* expressing premigratory neural crest cells in mutant embryos was indistinguishable from wild-type siblings up to approximately 18hpf (Figure 4.2G). In contrast to wild-types, 24hpf *cls* mutants exhibited a distinctly clustered pattern of neural crest cells in the head (Figure 4.3B). The number of *sox10*-positive cells around cranial ganglia and along cranial nerves was reduced (Figure 4.3B, D, F).

Furthermore, in *cls* mutants, neural crest cells appeared to accumulate in the premigratory area. We asked whether there were any differences in numbers of premigratory and migratory neural crest cells between *cls* mutants and wild-types. Thus, *sox10*-positive cells in the premigratory area, on the lateral and medial pathway were counted in the region of somites 1-5, 6-10, 11-15 and 16-20 of wild-type and $cls^{m618}$ mutant embryos at the 24, 30 and 35hpf stages (Figure 4.7). Generally, the number of *sox10*-positive cells on the medial and lateral pathway was reduced in more posterior regions at all stages, whereas numbers of premigratory neural crest cells increased. This is readily explained by caudal neural crest cells being a developmentally younger cell population and thus entering the migration pathways after more anteriorly located neural crest cells.

The mean number of *sox10*-positive premigratory neural crest cells per somite segment at 24hpf (9.4 ± 3.13) was broadly comparable to previous neural crest cell counts (11.3 ± 3.5; Raible, Wood et al. 1992). Thus, *sox10* might label all neural crest cells in the trunk at this stage.

As anticipated, the number of premigratory neural crest cells in wild-types decreased rapidly between the 24hpf and 30hpf stage as cells started migration and then remained fairly constant between the 30hpf and 35hpf stage (Figure 4.7). In contrast, the decrease in number of premigratory neural crest cells in mutant siblings was not as dramatic as in wild-types. This was observed as an accumulation of neural crest dorsal to the neural tube as a result of inhibited migration (Figure 4.3B, Q).

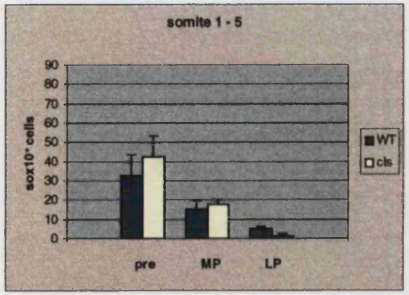
The number of *sox10*-positive cells on the medial pathway remained fairly constant at all stages investigated and no significant difference was observed between wild-type and *cls* mutants at any stage and in all regions (Figure 4.7).

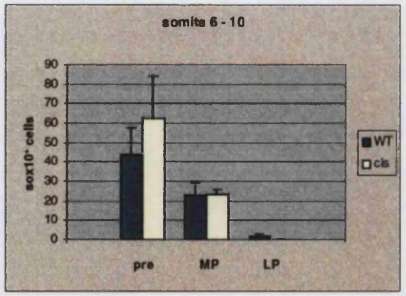
As described in section 4.1, pigment precursors rapidly down-regulated *sox10* and thus did not appear in counts of the 30 and 35hpf stage. Occasional *sox10*-positive cells were seen in migration on the lateral pathway, but this was only true in *cls* mutants and within somites 1-15 even in late stages (Appendix 4.1). Presumably, these represented neural crest precursors delayed in their differentiation, since equivalent pigment precursors in wild-types were differentiating by 30hpf and no longer expressed *sox10* (Figure 4.7).

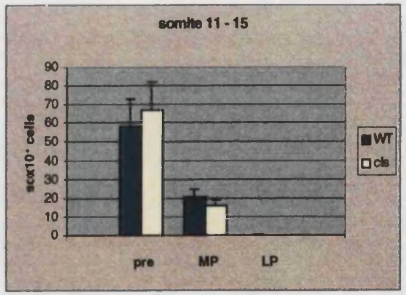
In summary, in *cls* mutants neural crest cells appeared to be inhibited in their migration on the lateral pathway. In contrast, *sox10*-positive cells on the medial pathway did not seem to be significantly affected (Figure 4.3B). By 36-40hpf, labelled cells were absent

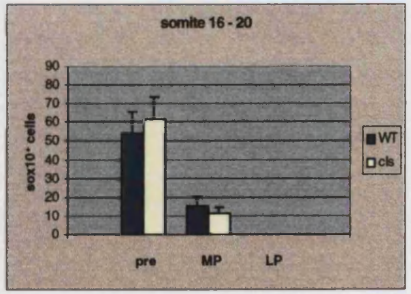
24hpf stage:

A



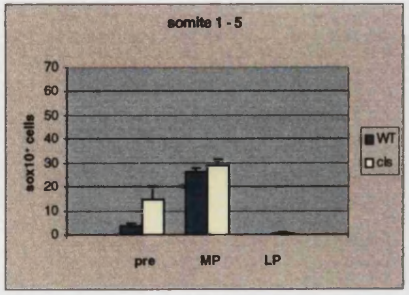



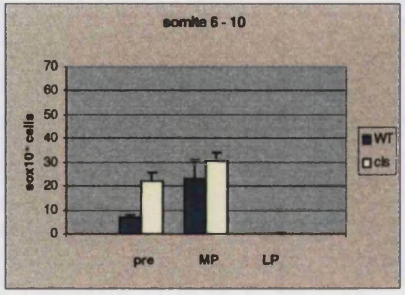



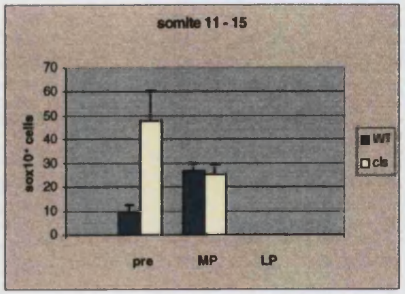



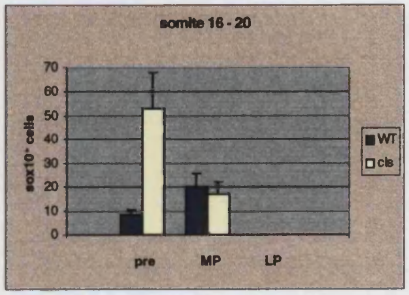

30hpf stage:

B



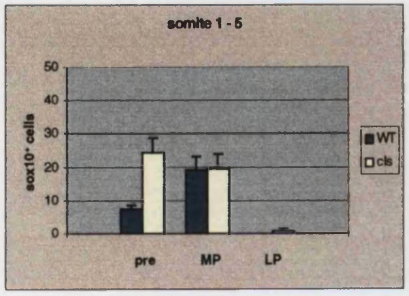



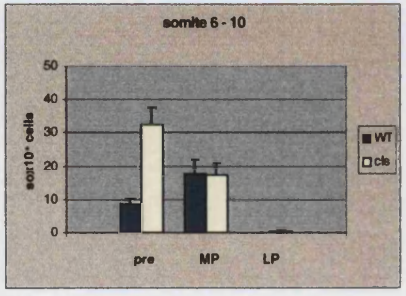



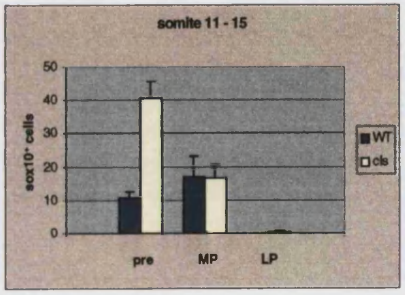



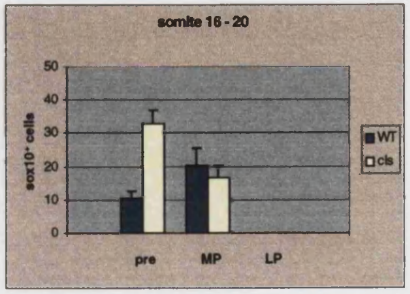

35hpf stage:

C



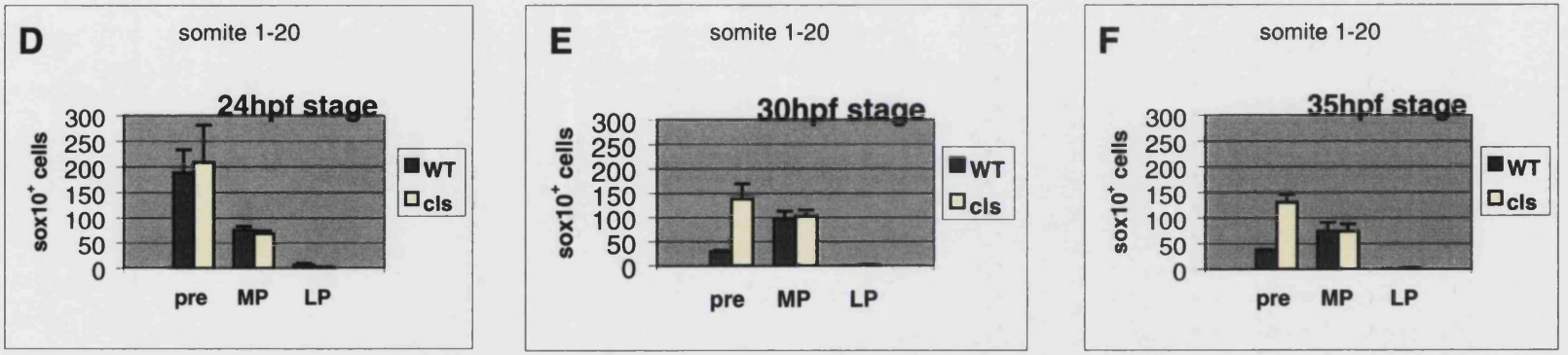



**Figure 4.7:** Location of *sox10* positive cells in wild-type and homozygous *cls*$^{m618}$ mutants

Cells expressing *sox10* were counted in 24hpf, 30hpf and 35hpf stage embryos from a heterozygous *cls*$^{m618}$ cross, with cells counted independently in the premigratory area (pre), on the medial (MP) and on the lateral migration pathways (LP) in each of regions corresponding to crest adjacent to somites 1-5, 6-10, 11-15 and 16-20 in wild-type (WT; blue speckled) and *cls* mutant (cls; yellow) embryos. Graphs present the mean number (error bars indicate standard deviation) of *sox10*-positive cells for each region and stage (n=5 for each genotype; A-C) or the mean total number of *sox10*-positive cells within somites 1-20 (D-F). In *cls* mutant embryos, *sox10*-positive neural crest cells tended to remain in the premigratory area at later stages, whereas they decreased rapidly in wild-types. Counts of cells on the medial pathway were broadly comparable between mutants and wild-type. Pigment cell precursors on the lateral pathway in wild-types rapidly turned off *sox10* expression as they started to differentiate and thus these cells no longer appear in counts of the 30hpf and 35hpf stages. By contrast, in *cls* mutants up to two *sox10*-positive cells were occasionally observed on the lateral pathway in more rostral somite segments at later stages.

from this latter location and only detectable on the dorsal neural tube and weakly around cranial ganglia in *cls* mutants (data not shown). By 60hpf, scattered cells, presumably satellite glia or undifferentiated neural crest cells (Britsch et al., 2001) associated with the posterior lateral line ganglion, showed weak expression (data not shown).
Enteric nervous system precursors expressing *sox10* were essentially absent in mutants by 60hpf (Figure 4.4K). Very rarely, one or two *sox10*-positive cells were observed in more proximal parts of the intestine (Figure 4.4L).

### *sox10 expression patterns in cls mutants outside the neural crest*

Outside the neural crest, the developing otic placode and vesicle in *cls* embryos strongly expressed *sox10* as described for wild-type siblings (Figure 4.2F, G, Figure 4.3B, D), although from approximately 40hpf onwards, the staining in mutant embryos decreased (data not shown).
No differences from wild-type siblings could be detected in the level, location and timing of expression in hindbrain (Figure 4.4K), spinal cord, jaw cartilage (Figure 4.4A, E) and pectoral fins (Figure 4.4E, I).

## Could Sox9a or Sox9b function redundantly with Sox10 in the neural crest?

*Sox9* homologues are the most closely related members of the *Sox* gene family to *Sox10*. Furthermore, just like mouse *Sox9* at E8.5, the zebrafish *sox9b* paralogue is expressed in the neural crest (Ng et al., 1997), whereas *sox9a* is only found in the lateral neural plate during early somitogenesis stages (E. Cheng, pers. commun.). Both paralogues were seen to be expressed in forming cartilage during later development (Cheng et al., 2000). Since ectomesenchymal derivatives such as craniofacial cartilage are unaffected in *cls* mutants, this could indicate some functional redundancy between Sox9 and Sox10 in the neural crest. We therefore asked, whether *sox9a* or *sox9b* expression co-localised with *sox10*-positive cells in the neural crest. To address this question, we performed double RNA *in situ* labelling with antisense probes against *sox10* and *sox9a* or *sox9b* on 5 somite and 14 somite stage embryos derived from a heterozygous $cls^{m618}$ cross. The 5 somite stage was chosen because a high percentage of cranial neural crest cells are fate-restricted and give rise to single cell-types (Schilling and Kimmel, 1994; Dorsky et al.,

1998). 14 somite stage embryos on the other hand are suitable since they allow comparisons between premigratory and migratory neural crest cell populations.

Double *in situ* hybridisation with *sox9a* and *sox10* revealed largely non-overlapping expression patterns at the 5 somite stage (Figure 4.8A-F). Although *sox9a* also labelled cells bordering the neural plate, they were found laterally to and at the posterior end of *sox10* expressing cells, possibly the region from which the otic placode arises. In fact, expressing regions appeared complementary. By 14 somites, both genes were expressed in the otic placode, but in only partially overlapping regions. *Sox9a* strongly labelled the centre of the placode (Figure 4.8 G, H) and in distinct patches on the dorsal rim (Figure 4.8I, J), whereas *sox10* was uniformly expressed in the entire otic epithelium. In contrast to *sox10*, *sox9a* was also seen in somitic mesoderm. Both genes were expressed in a patch just anterior to the otic placode, but the nature of these cells is unknown. They might be precursors contributing to the pharyngeal arches later during development. It was hard to eliminate the possibility that cranial neural crest cells expressed both markers. If *sox9a* transcripts were present, the signal was obscured by very strong *sox10* labelling. To investigate any possible crest expression of *sox9a*, this experiment has to be repeated with the blue *sox9a* signal developed more strongly than the red *sox10*. Alternatively, *sox9a* could be detected with the more sensitive fluorescent signal in a double or with the blue colour precipitate in a single *in situ* hybridisation on its own.

In contrast to *sox9a*, *sox9b* expression largely overlapped with *sox10* in the neural crest at all stages tested (Figure 4.9). However, in 5 somite stage embryos we observed neural crest regions that appeared to express predominantly either *sox9b* or *sox10*. Most $sox9b^{+}/sox10^{-}$ cells were found laterally at the rostralmost and caudalmost end of the neural crest region, whereas $sox9b^{-}/sox10^{+}$ cells seemed concentrated medially between areas of $sox9b^{+}/sox10^{-}$ expression (Figure 4.9B, D, F). Like *sox10*, *sox9b* expression was maintained in premigratory crest until the 14 somite stage and was also noted in the otic vesicle (Figure 4.9G, H, I). Caudalmost neural crest cells stained with *sox9b* alone (Figure 4.9A, B, H, I). This expression pattern could easily be explained by the earlier onset of *sox9b* transcription and thus *sox9b* expression reached more posteriorly than *sox10*. Differentially labelled cells were observed in the head reminiscent of *sox10* and *fkd6* double *in situ* hybridisations (Figure 4.9J, K, 4.7F). Like $fkd6^{+}/sox10^{-}$ cells, single $sox9b^{+}/sox10^{-}$ cells accumulated on the dorsal head just above the brain, posterior to the eyes, and few were seen dorsally between the otic placodes

**Figure 4.8:** Do *sox9a* and *sox10* expression domains overlap?
Double RNA *in situ* hybridisation with *sox9a* (brightfield, purple) and *sox10* (fluorescent, green) on 5 somite stage embryos indicates complementary expression domains.
(A-D) *sox10* (brightfield, red, A, C; fluorescent, green, B, D) is expressed in most neural crest cells (ncc), whereas *sox9a* (purple, A-D) appears to label a more lateral region from which the otic placode will arise. The region indicated by a black box in (C) enlarged in (E, F) again strongly indicates that patterns are largely complementary, although a possible overlap is seen in a narrow region (arrowhead). (G, H) In 14 somite stage embryos, *sox9a* (purple) is expressed in somites (s), in a patch in the centre of the otic placode (o, G, H and in dorsal view I, J) and a double labelled region just anterior to the otic placode (arrowhead in G-J). Any possible faint expression of *sox9a* in the neural crest was obscured by the strong *sox10* signal.
In lateral (A-B, E-H) and dorsal (C-D, I-J) views anterior is to the left. e, eye; ncc, neural crest cells; o, otic placode; s, somite. Scale bar: 155μm in A-D; 80μm in E-F; 115μm in G-H; 55μm in I-J.

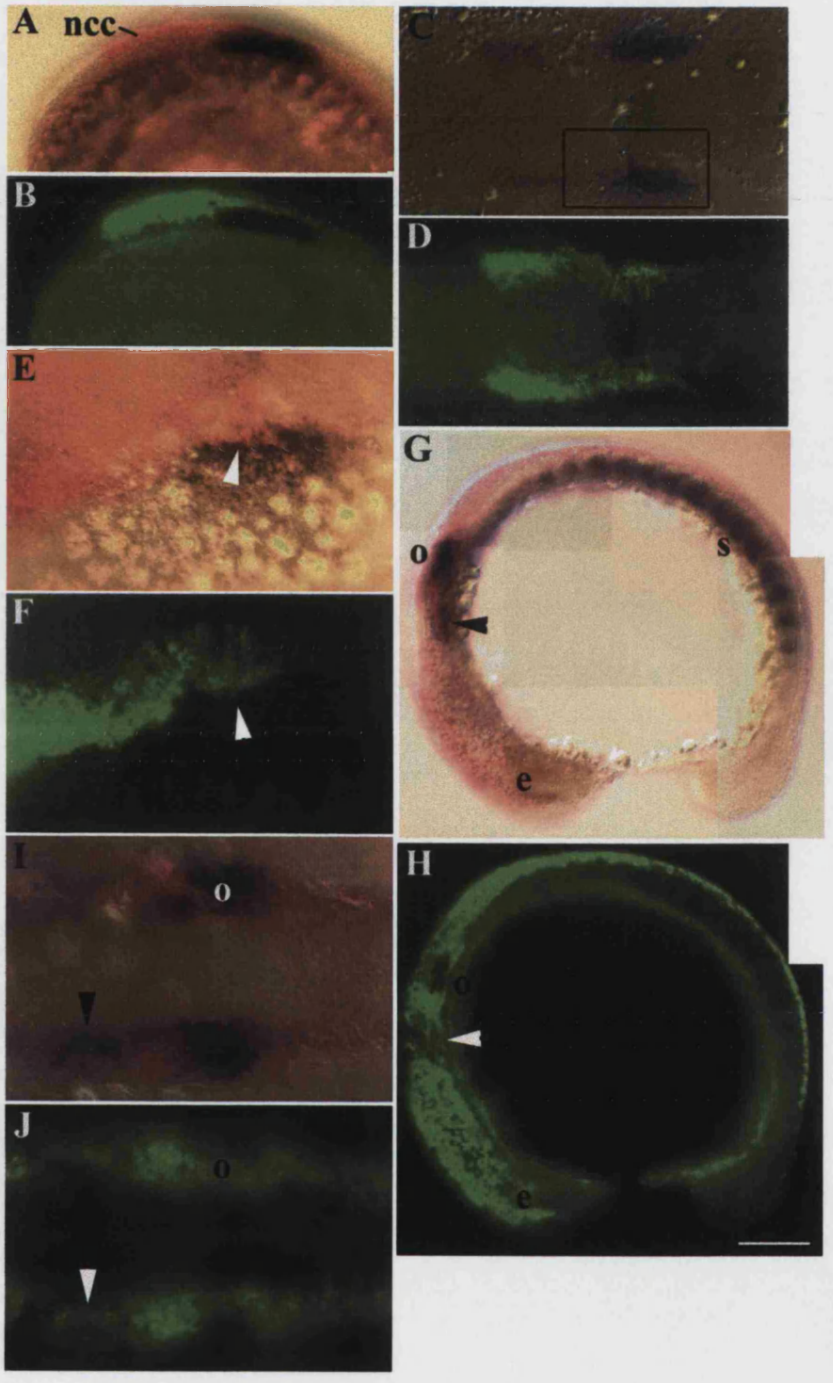
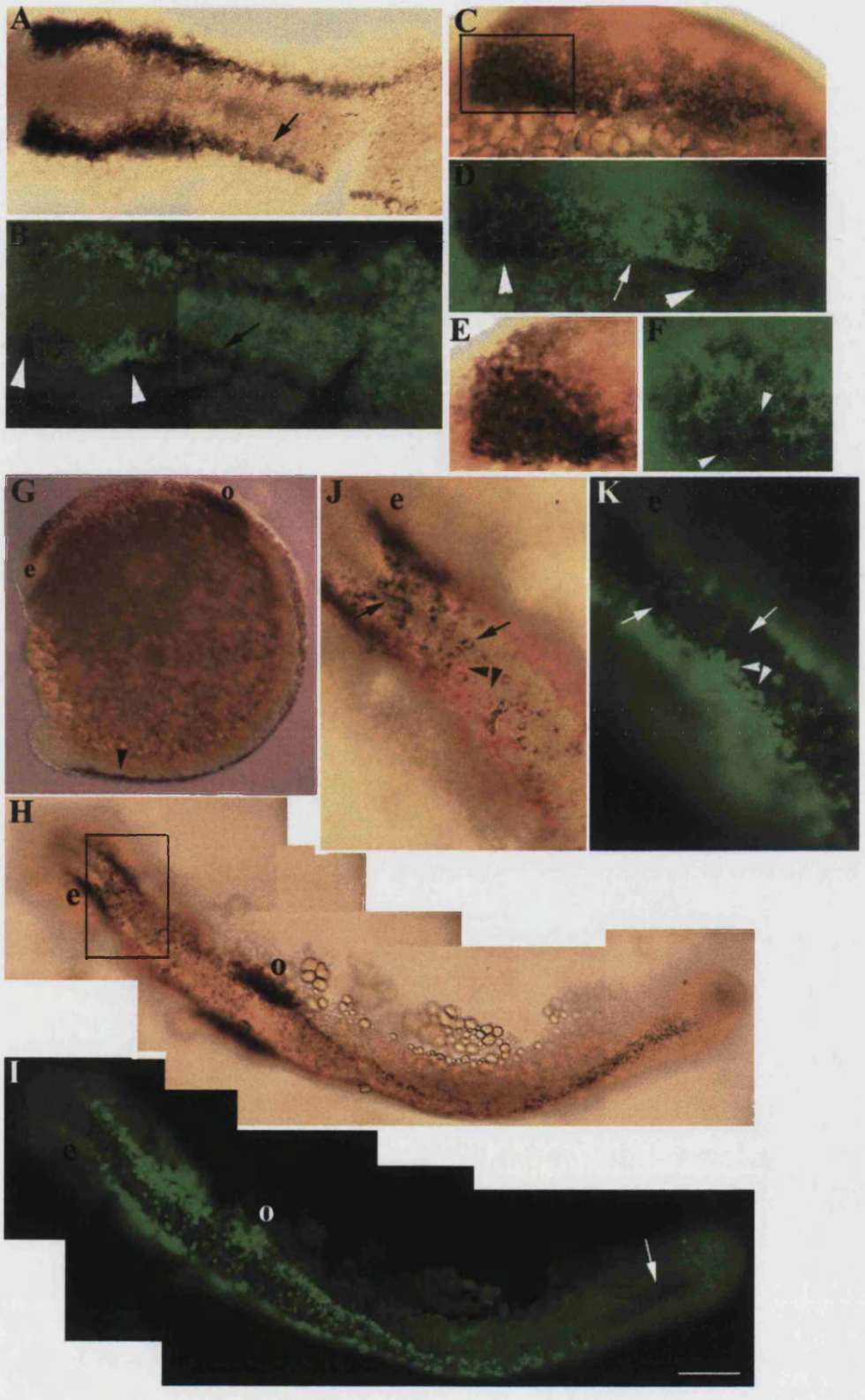

A
ncc
B
C
D
E
F
G
o
s
e
I
o
J
o
H
o
e

**Figure 4.9:** Do *sox9b* and *sox10* expression domains overlap?
Double RNA *in situ* hybridisation with *sox9b* (brightfield, purple) and *sox10* (brightfield, red; fluorescent, green) on 5 somite stage embryos shows a large overlap of expression in the neural crest. Areas of incomplete overlap are seen at the anterior border and lateral edges of the neural crest (white arrowheads, B, D, F) and in the posterior lateral neural plate (black arrow, A, B). The latter region is due to earlier onset of *sox9b* expression compared to *sox10*. (C, D) The largely double labelled neural crest region is shown in an enlarged lateral view (white arrowheads). The region in between appears to contain crest cells with strong *sox10*, but fairly weak *sox9b* expression (white arrow). The region outlined by a black box in (C) is shown in more detail in (E, F) to demonstrate clearly single *sox9b*$^{+}$/*sox10*$^{-}$ cells (white arrowheads). (G-I) A broad overlap of both markers in the neural crest and in the otic placode (o) is seen in 14 somite stage embryos. Again, the caudalmost neural crest only expresses *sox9b* (black arrowhead in G and white arrow in I). Particularly strong *sox9b* expression is seen around the eye (e). The region outlined by a black box (H, enlarged in J, K) shows single crest cells which express either *sox9b* (arrows) or *sox10* (arrowheads), located mostly above the brain, just posterior to the eye (e).
In dorsal (A-B, H-K) and lateral (C-F) views anterior is to the left, in the lateral view (G) anterior is up. e, eye; o, otic placode. Scale bar: 155μm in A-B; 85μm in C-D; 55μm in E-F, J-K; 130μm in G-I.


A
B
C
D
E
F
G
o
e
J
e
K
e
H
e
o
I
e
o

(Figure 4.9J, K). The staining of cells around the eye had a distinct dark purple hue, which suggests that both *sox9b* and *sox10* transcripts were present (Figure 4.9J, K). Whether each of these cells expressed one or the other or both markers could not be determined without sections. In contrast, areas of particularly strong *sox10* staining were seen laterally, between the eyes and the otic placode (Figure 4.9H, I). Scattered *sox9b*$^{-}$/*sox10*$^{+}$ cells were located in the same region on top of the neural tube (Figure 4.9J, K).

The expression of neither *sox9* paralogue showed any overt difference between *cls* and wild-type embryos at the 14 somite stage.

## *Discussion*

We have demonstrated that the expression pattern for the isolated zebrafish *sox10* was reminiscent of mammalian *Sox10* and is thus consistent with conservation of *Sox10* gene function. In wild-types, *sox10* expression was observed in premigratory and in some migratory neural crest cells. Once differentiated, *sox10* was only maintained in the enteric nervous system and glial cells of the peripheral nervous system (PNS) and in addition, it was expressed later in jaw cartilage. Furthermore, presumed oligodendrocytes in the central nervous system (CNS), the otic vesicle and cartilage cells in the pectoral fins expressed *sox10*.

### *sox10* in the premigratory neural crest

Premigratory neural crest cells expressed *sox10* from approximately the 1-2 somite stage, soon after transcriptional initiation of *fkd6*, a neural crest marker (Odenthal and Nuesslein-Volhard, 1998). The neural crest appeared to form normally in *sox10*$^{m618}$ (*cls*$^{m618}$) mutants and the *sox10* expression pattern was indistinguishable from wild-types until approximately 24hpf.

Very similar observations have been made in mouse, where *Sox10* is expressed in the premigratory crest of wild-types and homozygous *Sox10*$^{Dom/Dom}$ mouse mutants from E8.5. Only from E10.5, mutant expression patterns start to differ from wild-type siblings (Southard-Smith et al., 1998).

In addition, we were able to resolve the *sox10* expression pattern in premigratory crest in more detail than has previously been reported from studies in mouse. We asked

whether *sox10* was expressed in all neural crest cells and found that occasionally, single cranial neural crest cells appeared to lack *sox10* expression; this was confirmed by double *in situ* hybridisations with *sox10* and *fkd6* in 6 somite stage embryos. This result was intriguing, since premigratory cranial neural crest cells in zebrafish at this stage were shown to give rise to single cell fates only (Schilling and Kimmel, 1994). *cls* mutants lack overt defects in ectomesenchymal crest derivatives, and thus, these derivatives might not require Sox10 function. Unfortunately, there is no early marker for those derivatives available to determine the specification status of these *sox10*-negative cells.

Furthermore, we also demonstrated incomplete overlap of expression of *sox10* and *sox9b* in double *in situ* hybridisations. Since *Sox9* genes in mammals have roles in chondrogenesis, it would be interesting to perform a triple *in situ* hybridisation to determine whether these *sox10*$^{-}$/*fkd6*$^{+}$ were also *sox9b*$^{+}$.

However, it appeared that *sox10* labelled all premigratory trunk neural crest cells at 24hpf. Counts of *sox10*-positive cells in this region were comparable, but slightly lower than previous counts of premigratory neural crest cells (Raible et al., 1992). One explanation why the cell number of previous cell counts was slightly higher may be due to different observers and perhaps small differences in the stage at which counts were carried out. As demonstrated in counts of premigratory neural crest cells at later stages, the number of *sox10*-positive cells in this location decreases over time and thus, the stage at which cell counts are carried out is important, especially after 24hpf.

Thus, in agreement with mammalian data, *sox10* is expressed and maintained in premigratory neural crest. However, *sox10* was found to be transcribed in only a large subset of precursors of unknown specification, at least in the cranial neural crest. Thus, use of *Sox10* expression as a marker for multipotent neural crest may be misleading (e.g. Pattyn et al., 1999).

## *sox10* in migratory neural crest cells and differentiating pigment cells

Neural crest cells on both migratory pathways in wild-types maintained *sox10* expression. Although neural crest cells have entered the lateral pathway by 24hpf (Raible et al., 1992), the number of *sox10*-positive cells on this pathway, which exclusively gives rise to pigment cell-types (Raible and Eisen, 1994), decreased rapidly

after 30hpf. This was explained by a downregulation of *sox10* as melanoblasts started to differentiate. At the same time, melanised cells, most of which appeared to lack *sox10* expression, were migrating on the medial pathway.
This was in agreement with reports in mouse, in which extensive *Sox10* expression was noted on the medial pathway and in presumptive melanoblasts in a region consistent with the dorsolateral pathway (Southard-Smith et al., 1998). However, it is unclear whether *Sox10* was also rapidly turned off in differentiated melanocytes.

In *cls* mutants, the initial migration on the medial pathway appeared normal at 24hpf. Presumably, these cells contributed to DRG neurons, cranial and DRG glial cells, which are less affected in *cls* embryos. Anterior DRG in *cls* mutants are of approximately normal size, whereas they are missing in far posterior segments (Kelsh and Eisen, 2000b). However, from 24hpf, neural crest cells seemed to accumulate in their premigratory position. Consistent with these observations, we could demonstrate in comparative counts that *cls* mutants had more *sox10*-positive cells in the premigratory area than wild-type siblings of the same stage. In wild-types, the number of premigratory neural crest cells rapidly decreased in older stages as expected, whereas in mutants the decrease was minimal up to 35hpf. In addition, hardly any *sox10*-positive cells or differentiated pigment cells were seen on the lateral pathway in mutants. Thus, our observations were consistent with previous proposals that the migration of neural crest cells in *cls* mutants was inhibited, especially on the lateral pathway (Kelsh et al., 1996; Kelsh and Eisen, 2000b). The clustering of neural crest cells observed in the head of 24hpf *cls* mutant embryos might represent yet another example of this migrational defect. These cells are scattered in wild-type heads presumably due to active migration. Similar results to our observations were reported in mice. Initial neural crest migration also appeared to be normal in mice homozygous for a targeted *Sox10* deletion until E10.5 (Britsch et al., 2001). However, in homo- and heterozygous $Sox10^{Dom}$ mutants, migrating neural crest cells labelled with *Ednrb* were developmentally delayed in their pathway from E10.5 and melanoblasts were absent in homozygotes (Southard-Smith et al., 1998). Detailed cell counts analysing the number of premigratory and migratory *sox10*-positive cells in wild-types and mutants of various stages have not been reported in mouse.

## *sox10* in peripheral and cranial ganglia neurons

Strong *sox10* expression was associated with forming cranial ganglia at 24hpf in wild-types. In contrast to cranial ganglia neurons, satellite glial cells surrounding cranial ganglia maintained high *sox10* expression levels in wild-types.

From 24hpf, *cls* mutants were distinguishable from wild-types by a reduction of *sox10* expression in satellite glia and Schwann cells along cranial neurons.

Comparable results were obtained in mouse and chicken. *Sox10* staining was associated with cranial ganglia, but hybridisation signals for *Sox10* and the neuronal marker SorLA in mouse were non-overlapping in the developing trigeminal ganglion (Kuhlbrodt et al., 1998a). This was also demonstrated for neural crest derived ganglia in the head and trunk of chicken (Cheng et al., 2000). Although indistinguishable from wild-type up to E9.5, by E10.5 $Sox10^{Dom/Dom}$ mice showed strongly reduced levels of *Sox10* expression associated with cranial ganglia (Southard-Smith et al., 1998).

Expression of *sox10* associated with DRG was difficult to distinguish from other *sox10*-positive cells on the medial pathway at the stages examined. Even at later stages, double labelling with *sox10* and anti-Hu antibody will be necessary to determine whether *sox10* is expressed in DRG neurons or glia.

In rodents and chicken, *Sox10* expression was observed in DRG, but the exact neural cell-type was not determined (Kuhlbrodt et al., 1998a; Southard-Smith et al., 1998; Cheng et al., 2000).

## *sox10* in peripheral glial cells

Strong *sox10* expression was maintained in cranial ganglia glia and Schwann cells associated with the posterior lateral line nerve and other cranial and sensory nerves at least up to 60hpf in wild-types. Thus, the *sox10* expression pattern in those locations was reminiscent of the glial marker *fkd6* at the stages examined (Kelsh et al., 2000a) and did not overlap with the neuronal marker anti-Hu antibody. In *cls* mutants, *sox10* expression associated with all peripheral ganglia, presumably glial cells, was reduced from 24hpf.

Studies in rodents also reported extensive *Sox10* expression in cells associated with nerve fibres, consistent with them representing Schwann cells. In particular, the trigeminal nerve, the sympathetic trunk and even on motor nerves, which originate from the CNS, not ganglia are described (Kuhlbrodt et al., 1998a). In chicken, sections

through DRG double labelled with *Sox10* and anti-Hu antibody also showed a non-overlap of *Sox10*-positive cells with neuronal derivatives (Cheng et al., 2000). As an alternative explanation for these *Sox10*-positive cells, undifferentiated neural crest cells have been suggested in mice, since they did not express neuronal or glial markers (Britsch et al., 2001). In $Sox10^{Dom/Dom}$ mutants, expression of *Sox10* in cranial ganglia was drastically reduced by E10.5 and absent from E11.5 (SouthardSmith et al., 1998). Thus, *sox10* expression in peripheral glial cells seemed to be conserved between vertebrates and fish. Since mice lack the posterior lateral line nerve, zebrafish show this additional region of *sox10* expression. The suggested presence of undifferentiated neural crest cells around cranial ganglia has not been investigated in zebrafish. However, segmentally arranged *sox10*-positive cells at 35hpf in the trunk were thought to be glial cells associated with the DRG or Schwann cells associated with motorneurons. Since they only expressed *sox10* up to 36-40hpf, in contrast to other glial derivatives, they too might be undifferentiated neural crest cells instead.

## *sox10* in the enteric nervous system

Enteric ganglia precursors were labelled by *sox10* from 24hpf (R Kelsh, unpubl. observ.). Numerous *sox10*-positive cells were seen in sections of wild-type guts surrounding the intestinal lumen at 60hpf, but these cells were missing in *cls* mutant embryos. However occasionally, a few *sox10* labelled enteric precursors were also observed in mutants.

The results obtained in zebrafish paralleled those in mouse. *Sox10* transcripts detectable at E10.5 in wild-types, were absent in homozygous $Sox10^{lacZ}/Sox10^{lacZ}$ knock-in mice (Southard-Smith et al., 1998; Britsch et al., 2001). Furthermore, the degree of intestinal innervation also showed a range from severe hypoganglionosis to aganglionosis (Puliti et al., 1996). These differences were explained by the presence of genetic modifier loci (Kapur et al., 1996).

## *sox10* in skeletal neural crest derivatives

A very interesting result represented the staining seen in the developing jaw cartilage. This was the first demonstration of *sox10* expression in this tissue. Although craniofacial cartilage precursors are labelled by *dlx2* at 29hpf, they did not express

*sox10* until 48hpf. By this time, *sox10* expressing chondrocytes are arranged in a "stack-of-coins" structure. Consistent with our observations and onset of *sox10* expression, facial cartilage condenses and starts to differentiate between 2 and 3 days of development (Kimmel et al., 1995). This might indicate a role for Sox10 in these cartilages. Although *cls* mutants lack overt defects in the craniofacial skeleton, late mutants showed a slight retardation of arch development (Kelsh and Eisen, 2000b).
The median fin mesenchyme, another ectomesenchymal neural crest derivative, showed no expression of *sox10* at any stage tested.
*Sox10* in mouse or chicken has never been observed in craniofacial cartilage precursors. However, this might be another example of the cell-biological advantages of zebrafish due to its transparent embryos allowing higher sensitivity in *in situ* hybridisations. A second example of previously undetected *Sox10* expression in mesenchymal condensations in mouse will be discussed below. Our result might be comparable to staining reported in the cephalic crest, which also gives rise to nasal bones in 6 week old human embryos (Bondurand et al., 1998a).

## *sox10* in the pectoral fin buds and pineal gland

*sox10* expression in the developing pectoral fins seemed to be the first report of *sox10* in this tissue. No staining was reported in functionally equivalent expression domains in chicken (Cheng et al., 2000). Only very recently, weak β-galactosidase staining was discovered in mesenchymal condensations of the digits, radius and ulna in $Sox10^{lacZ}$ knock-in mice (Britsch et al., 2001). Presumably, due to low levels of expression, these structures had been overlooked in less sensitive *in situ* hybridisations.
The nature of these *sox10*-positive cells in zebrafish has not been further investigated, but consistent with their location and "stack-of coins" arrangement, they too are likely to be differentiating non-neural crest derived cartilage. Labelled cells were not observed in the apical ectodermal fin folds, similar to a pattern seen in *sox9a* and *sox9b* labelled pectoral fins at 60hpf (data not shown; Chiang et al., 2001). Furthermore, the pectoral fin expression of *sox10* in *cls* mutants was indistinguishable from wild-types, just as no overt differences in labelling intensities were noted between homo- and heterozygous $Sox10^{lacZ}$ mice (Britsch et al., 2001).

At 24hpf, the developing pineal organ expressed *sox10*. Apart from zebrafish, *sox10* expression in this tissue has only previously been observed in chicken (Cheng et

al., 2000). This gland arises through delamination of cells from the neuroepithelium similarly to neural crest cells and thus Cheng and coworkers proposed a role for Sox10 in enabling this process by altering the expression of adhesion molecules. This area of *sox10* expression intriguingly coincides with *sox9a*, *sox9b* (Chiang et al., 2001) and *fkd6* expression. The roles and possible interactions of the above transcriptions factors in pineal development are still unknown.

## *sox10* in the central nervous system (CNS)

In zebrafish, expression in the developing CNS was noted from at least 24hpf up to approximately 48hpf in the ventricular zone. In chicken, this region gives rise to neuronal and glial precursors (Cheng et al., 2000). Consistent with this, we observed *sox10* expression in putative oligodendrocytes from approximately 48hpf up to at least 60hpf.

Thus far, it has not been demonstrated whether *Sox10* in mice is maintained continuously in glial cells until adulthood, where it is found mostly in oligodendrocytes in the adult CNS (Kuhlbrodt et al., 1998a). In the PNS, *Sox10* expression associated with peripheral ganglia and nerves was maintained up to P21 and along the sciatic nerve in adults (Kuhlbrodt et al., 1998a). Alternatively, transcription could be regulated independently in embryos and adults. An interesting future line of research to address this question would extend these expression studies to later stage zebrafish.

As an example to demonstrate the importance of SOX10 during postnatal human development, a *SOX10* mutation in an 11 year old patient has been associated with deficiencies in myelin, produced by oligodendrocytes and Schwann cells in the CNS and PNS, respectively. The patient suffered from seizures, sensory dullness and reduced nerve conduction velocities in addition to the classical symptoms of WS4 (Inoue et al., 1999).

## *sox10* in the developing ear

The onset of *sox10* expression in the lateral head correlated approximately with the first appearance of the otic placode at around the 10 somite stage (14hpf; Whitfield et al., 1996) and strong labelling of the otic epithelium was maintained at all stages examined, at least up to 60hpf. This is similar to expression data in chicken, where the first

expression of *Sox10* is seen in the otic placode at HH stage 9 (~6-7 somites; Cheng et al., 2000). Expression in the developing ear was maintained at least until HH stage 26 in chicken (Cheng et al., 2000) and E16.5 and adults in mouse (Watanabe et al., 2000). *cls* embryos show defects in ear morphology and thus suggest a role for Sox10 in otic development. The first visible defect in the otic vesicle of *cls* mutant embryos has been noted at approximately 36hpf (Malicki et al., 1996) and by 48hpf, only 2 instead of 3 cristae, labelled with *mshC*, were observed (Whitfield et al., 1996). Consistent with these observations, the labelling intensity of *sox10* in the *cls* otic epithelium started to decrease at around the same time, 40hpf. By 5 dpf, *cls* mutant ears tended to be smaller than wild-types, with two small otoliths and a single sensory patch (Whitfield et al., 1996). Thus, these sensory and morphogenetic defects strongly suggest a dysfunctional inner ear. This is consistent with human *SOX10* mutations, even though the latter is usually attributed to a loss of melanocytes in the stria vascularis of the cochlea (Steel, 1989; Pingault et al., 1998a) not present in fish.
*Sox9a* and *sox9b* were also expressed in the otic placode to various extents. It will certainly be an interesting line of future research to investigate the roles and possible interactions of these 3 transcription factors in this organ.

In summary, we provided evidence that the cloned *sox10* orthologue shows conservation of its expression patterns with mammalian *Sox10s*. We showed that *sox10* was strongly expressed in neural crest derivatives, consistent with our proposal of *sox10* as a candidate gene for the *cls* mutation. Furthermore, we demonstrated that *sox10* expression was disrupted in *cls* mutant embryos. In the next chapter we will investigate whether *sox10* and *cls* are tightly linked, as predicted if *cls* encodes *sox10*.

# Chapter 5 – Mapping of *sox10* on the radiation hybrid panel LN54

## *Introduction*

In chapters 3 and 4 we have described the isolation and expression pattern of a *sox10*-like homologue. The results were consistent with this *sox10*-like gene being a true *sox10* homologue and a strong candidate for the gene encoded by the *cls* locus. To further test this hypothesis, we asked whether the *cls* mutant locus maps to the same region on one of the linkage groups as *sox10*.

The $1.7x10^9$ base pair zebrafish genome, approximately half the size of the mammalian genome, is contained within 25 linkage groups (LG) each corresponding to a chromosome. Each linkage group has been characterised by polymorphic markers such as RAPDs (Random Amplified Polymorphic DNAs), SSLPs (Simple Sequence-Length Polymorphisms) and RFLPs (Restriction Fragment Length Polymorphisms) that provide a framework into which genes and mutations of interest can be placed.

## The *cls* mutation is located on the distal end of LG 3

To test whether *cls* and *sox10* map to the same region, the mutation had to be placed on the linkage map. During a short collaborational visit to the Max Planck Institute für Entwicklungsbiologie in Tübingen, Germany, Susana Lopes mapped *cls* to the distal end of LG 3 (Dutton et al., 2001b).

To do so, Susana generated a mapping cross between $cls^{tw1}$, bred onto a Tuebingen wild-type background, and WIK wild-type zebrafish. These two different wild-type backgrounds are polymorphic for many microsatellite markers necessary to establish linkage of the *cls* locus.

First she asked on which of the 25 linkage groups *cls* was located. To test this, microsatellite markers on each linkage group polymorphic in those two backgrounds were assessed for linkage to the *cls* mutant locus by PCR. If a marker was located on a different linkage group to *cls*, recombination between those loci would be observed very frequently. On the other hand, the tighter the linkage and thus the closer marker and mutant locus are positioned to each other, the lower the probability of recombination in the sequence between them. Hence, the distance between markers is measured in

centiMorgans (cM), where 1cM represents 1% recombinants. Based on the relation of the physical (~$1.7x10^9$ bp) and genetic size (~2720cM) of the zebrafish genome, 1cM is estimated to be equivalent to approximately 600kb (Postlethwait et al., 1994). Recombination between *cls* and markers on LG 3 were less frequent than with markers for any other linkage group. Thus, *cls* was tentatively placed on LG 3.
Next, she asked which region of LG 3 contained the *cls* mutant locus. This was achieved by testing recombination rates between different markers on LG 3 and *cls*. Again, the marker showing the least recombination was expected to be located nearest to the mutant locus. Thus, fine mapping of 274 meioses finally positioned *cls* within a 3.9 cM (2.3Mb) interval between markers z872 and z13387 on the distal end of LG 3 (Figure 5.1).

If *cls* encoded *sox10,* the gene would be expected to map to the same region on LG 3. There are several methods to map a cloned gene. The location of a gene can be determined by fluorescent *in situ* hybridisation (FISH), by linkage mapping or by radiation hybrid mapping. Syntenic regions can also provide a putative location, if homologues of the cloned gene have been identified and mapped in other species.
The FISH technique has been established in zebrafish, but is technically challenging (Amores and Postlethwait, 1999). There are syntenic regions found on the chromosome containing *Sox10/SOX10* between mouse chromosome 15, human chromosome 22q and zebrafish LG 3 and LG 12. However, these regions don't include the locus itself. Thus, they only provide an indication that *sox10* might be located on the same LG as *cls*.
Classical linkage mapping requires the identification of a polymorphism in the cloned gene between different strains such as a restriction fragment length polymorphism (RFLP). However, this is not required for radiation hybrid mapping making it a straightforward method to use. Furthermore, there was the LN54 radiation hybrid panel (Hukriede et al., 1999) available to our lab, which had been used successfully to map genes. Thus, mapping *sox10* on this panel appeared to be the most suitable approach.

## The radiation hybrid panel LN54

Radiation hybrid panels represent a very convenient method of mapping a gene to a chromosomal region, since there is no requirement to identify polymorphisms. The radiation hybrid panel LN54 (Loeb, NIH, 5000 rad, 4000rad) was generated by Mark Ekker and his colleagues in the Loeb Research Institute, Ottawa, Canada in 1999

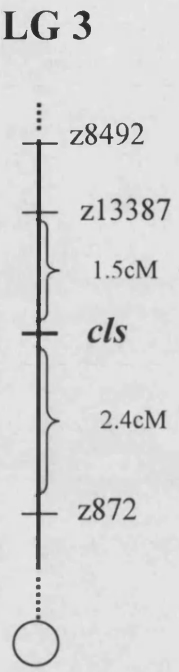



**Figure 5.1:** The *cls* mutation is located on the distal end of Linkage group 3 (LG 3) Linkage mapping of 274 meioses (carried out by Susana Lopes) placed *cls* in a 3.9cM interval between microsatellite markers z13387 and z872 (Dutton et al., 2001). 1cM (centiMorgan) represents 1% recombinants or approximately 600kb. The proximal centromeric end is indicated by the ◯ .

(Hukriede et al., 1999).

The panel represents a collection of genomic DNA isolated from 93 mouse cell lines each of which have incorporated various amounts of fragmented zebrafish genomic DNA. Each of the hybrid cell lines has been frame-work mapped with microsatellite markers specific for each zebrafish linkage group. The cell lines which contain the gene to be mapped (positive hybrids) are identified by PCR amplification of a gene specific product. The chromosomal region(s) common to all positive hybrids contains the gene in question with a certain probability (lod score). The lod score is defined as the log10 of the ratio of the probability that two markers are linked with a given recombination value and the probability that they are not linked. (Hukriede et al., 1999).

To generate the radiation hybrid panel, zebrafish fin AB9 donor cells were irradiated to introduce random chromosomal breaks. These irradiated cells were then fused to mouse B78 melanoma recipient cells. The resulting radiation hybrid cell lines retained various amounts of zebrafish chromosomal fragments. The percentage of the zebrafish to mouse chromosomal DNA ratio is called the retention rate. For selection, zebrafish chromosomes had been tagged with the aminoglycoside phosphotransferase (*neo*) gene, which rendered successfully fused hybrids resistant to gentamicin (G418). Out of several radiation doses tested, 5000rad and 4000rad, hence the name of the panel, produced hybrids with suitable retention rates of 18-21%. In total, 93 lines, 81 from the 5000rad and 12 from the 4000rad irradiation dose, were expanded with an overall retention rate of 22%. Thus, a hybrid collection was created, large enough to be used for over 75000 assays in duplicate. The panel was analysed for completeness by characterising it with a total of 1053 markers, including SSLPs, ESTs, and sequence-tagged sites (STS) of cloned genes. The probability of mapping a marker or a gene on this panel was found to be 88%, which serves as an estimation for genome coverage. The distance between markers is measured in centiRays (cR). 1cR represents 1% frequency of a breakage occurring between two markers after exposure to a specific radiation dose. With the radiation used to generate the LN54 panel, the average breakage frequency was estimated to be 148kb per cR. The resolution of the map corresponds to the average interval size between markers used to create the framework. With 1cR=148kb the resolution is estimated to be approximately 500kb. After characterisation of the 93 hybrid lines, 4 possible gaps were identified, which might just represent regions where the marker density was too low to be mapped adequately.

However, over time and with the addition of newly isolated genes placed on the panel, the coverage is expected to increase.

## *Results*

### Mapping *sox10* on the radiation hybrid panel LN54

Mapping of a gene on the radiation hybrid panel involves a PCR based screen. Gene specific primers are used to amplify a gene specific fragment from DNA isolated from each hybrid cell line. Those that contain the *sox10* genomic region will show amplification of this gene specific product. Thus, depending on the combination of hybrid cell lines positive in this PCR assay, a likely position of this gene can be deduced.

Thus, to map *sox10* on the radiation hybrid panel, two gene specific primers S11 and S13 were designed (Materials and Methods). They are located within the 3'UTR of *sox10* and amplify a 931bp fragment (Figure 5.2). This fragment is highly specific to *sox10* since there is very little sequence conservation outside the coding region.
The PCR conditions for this primer pair were optimised prior to the mapping assay. Conditions had to be obtained that reliably amplified a specific product from 100ng zebrafish genomic DNA and from a 1:10 mixture prepared from zebrafish and mouse genomic DNA, but not from 100ng of mouse genomic DNA alone. This was important to ensure no false positives or negatives were scored. The PCR conditions and the PCR program are described in Materials and Methods.
PCR reactions were set up with 100ng of genomic DNA from each of the 93 radiation hybrid cell lines and with 3 control DNAs; the donor zebrafish cell line AB9, the recipient B78 mouse cells and a 1:10 mixture of both (Figure 5.3). After amplification the reactions were then scored on a standard 1% agarose gel for specific products. To ensure the reliability of the mapping data obtained the PCR assay must be performed at least in duplicate, and in triplicate if more than 5% differences between the first two assays are experienced.
The first two screens that were undertaken only produced very faint products which were difficult to score reliably. This might be explained by differences in the genomic DNA samples. The genomic DNA that was used to optimise the PCR conditions had been prepared in our laboratory, whereas hybrid DNAs had been kindly provided by

# Location of primers for mapping *sox10* on the LN54 radiation hybrid panel

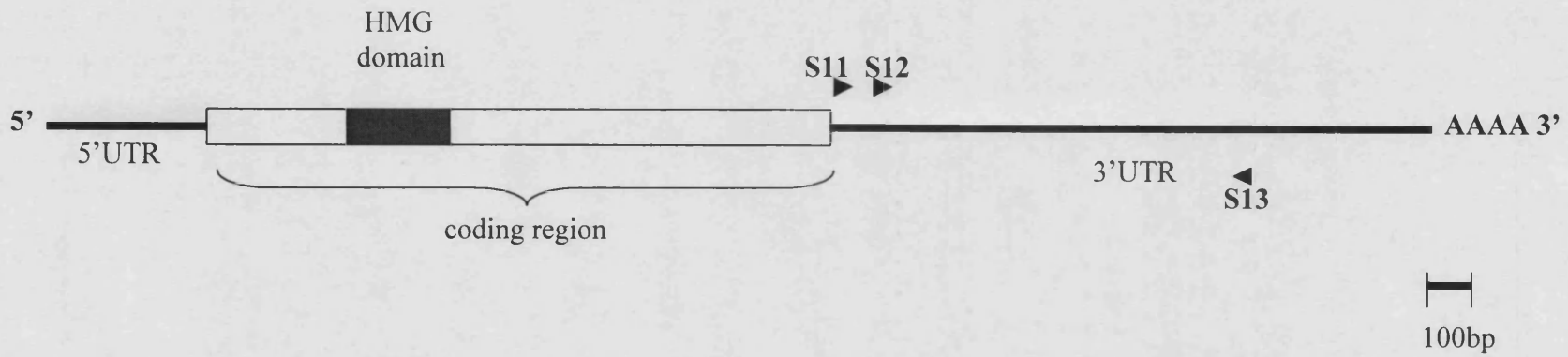



**Figure 5.2:** The location of primers S11, S12 and S13 within the *sox10* 3' UTR

The schematic shows the *sox10* cDNA including coding and untranslated regions.

Mapping assays were performed by amplification of a *sox10* specific 931bp product between primers S11 and S13. Primers were designed to the 3' untranslated region (3'UTR) to ensure specificity to sox10 (non-coding regions show very low sequence conservation). Radiation hybrids scoring positive in the screen retained a fragment of the linkage group containing *sox10*. To increase the amount of specific product, nested PCRs were set up with primers S12 and S13. The scale bar indicates the length of a 100bp fragment.

| 1 | 2 | 4 | 5 | 6 | 8 | 9 | 11 |
|---|---|---|---|---|---|---|---|
| 13 | 16 | 18 | 19 | 23 | 25 | 27 | 29 |
| 30 | 33 | 36 | 39 | 40 | 41 | 47 | 48 |
| 49 | 50 | 52 | 54 | 57 | 59 | 63 | 65 |
| 66 | 70 | 73 | 74 | 79 | 80 | 83 | 84 |
| 85 | 86 | 87 | 88 | 89 | 991 | 92 | 96 |
| 97 | 98 | 101 | 104 | 105 | 106 | 108 | 109 |
| 114 | 117 | 119 | 121 | 123 | 125 | 132 | 135 |
| 136 | 137 | 138 | 150 | 151 | 152 | 153 | 167 |
| 169 | 174 | 175 | 176 | 178 | 182 | 183 | 184 |
| 190 | 300 | 301 | 302 | 303 | 304 | 305 | 306 |
| 308 | 309 | 310 | 311 | 312 | mix | AB9 | B78 |

**Figure 5.3:** The LN54 radiation hybrid panel

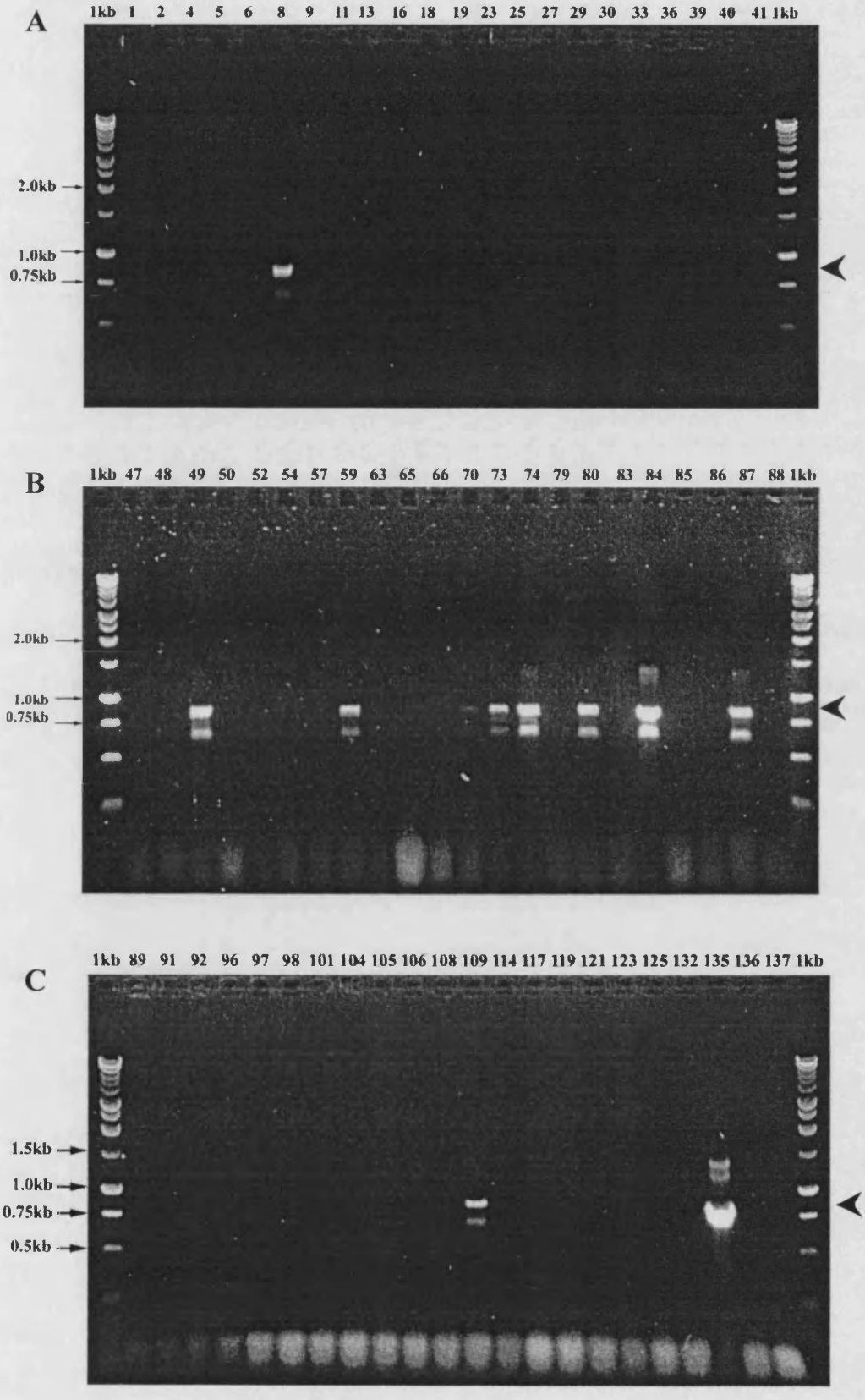
A total of 93 hybrid cell lines labelled 1-312, 81 from the 5000rad and 12 from the 4000rad irradiation dose, had been chosen and DNA included in the panel (Hukriede et al., 1999). The sample labelled "mix" represents $1/10^{th}$ zebrafish and $9/10^{th}$ mouse genomic DNA, whereas "AB9" contains genomic DNA from the zebrafish donor cell line and "B78" genomic DNA from the mouse recipient cell line. All three represent control DNA pools which are included in each mapping assay.

Ekker. It was decided to increase the amount of product by either subjecting the PCR reactions to more cycles of amplification or by nested PCR. In a pilot study with a mixture of 1:10 zebrafish to mouse genomic DNA mix, increasing the number of cycles did not improve the screen result, whereas setting up a nested PCR reaction with nested primers S12 and S13 (Figure 5.2) proved to be very successful. Thus, after the first round of PCR with primers S11 and S13, the product was diluted 1:10 in milli-Q water and subjected to an equivalent reaction setup and PCR program as before, but including primers S12 and S13. The entire PCR reaction was loaded onto a standard agarose gel and scored for presence or absence of the specific product with the size of 863bp. Radiation hybrids 8, 49, 59, 70, 73, 74, 80, 84, 87, 109, 135, 190, 304, 306, 308 and 311 proved positive. Usually, two bright bands were observed. One of the expected size and one approximately 150bp smaller (Figure 5.4). As anticipated, no bands were observed in the control reaction containing mouse genomic DNA, but high levels of product were derived from zebrafish genomic DNA and also from the 1:10 mix of zebrafish and mouse genomic DNA. Individual radiation hybrids that proved positive in this first assay were recorded and the whole assay repeated a second time. This duplicate assay did not work as well as the first attempt and gave slightly weaker products, most likely because a fifth of the Taq DNA polymerase from Advanced Biotechnologies was replaced by Taq DNA polymerase from Promega. In this duplicate assay products were observed in 15 out of 16 (93.8%) hybrid cell lines previously positive in the first assay. Hybrid number 70, which produced the faintest band in the first assay, did not give a detectable product in the second assay. All previously negative hybrid cell lines were consistently negative in the duplicate screen. Since only one hybrid cell line gave an ambiguous result, the data was considered robust enough for submission.

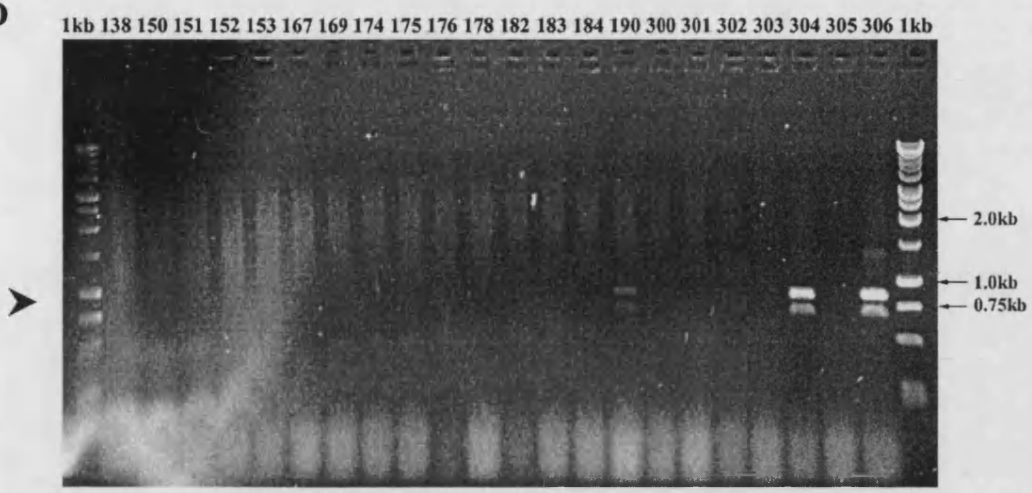
Mapping results are returned as two lod scores indicating the probability of linkage to the two most likely markers on the 25 linkage groups. The lod score is defined as the logarithm of the likelihood ratio for linkage (Hukriede et al., 1999). The best lod score must be greater than a lod of five in order to be considered significantly linked. Also, if the difference between the best lod score and the lod score on the second best LG is less than three, this could be indicative of multiple linkage.

The mapping data for *sox10* returned a best lod score of 17.6 for the marker z8492, which is located on LG 3, 0cR from *sox10* (Figure 5.5). The best marker in the second best linkage group, LG 1, is linked to *sox10* with a lod score of 5.4. In Figure 5.5, markers used in linkage mapping of the *cls* locus and framework markers derived from

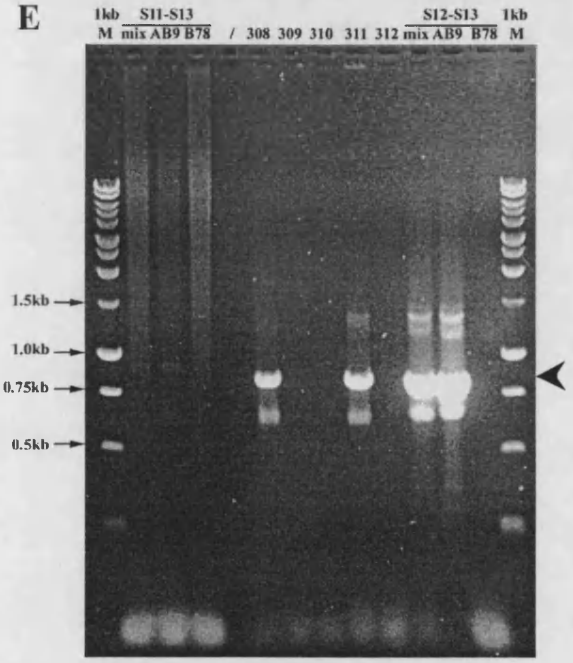
**Figure 5.4:** Mapping *sox10* on the radiation hybrid panel LN54.
The first round PCR between primers S11 and S13 only produced a few very faint products (data not shown). The control reaction of this PCR are shown in (E). In the PCR reaction labelled "S11-S13 mix" containing zebrafish and mouse genomic DNA in a 1:10 ratio, no product was detectable after amplification with S11 and S13. However, a strong product of the correct size (863bp, arrowhead) was observed after nested PCR with S12 and S13 (lane S12-S13 mix). As expected, *sox10* specific bands were amplified from zebrafish genomic DNA (lanes AB9) in both rounds of PCR and no product was ever detectable using mouse genomic DNA as a template (lanes B78). No sample was loaded in the lane marked with a dash, /. Thus, nested PCR reactions in duplicate were used to map sox10 on the LN54 panel. Panels (A-E) show examples of the nested amplification between primers S12 and S13 from 93 radiation hybrids (numbers 1-312), controls (lanes mix, AB9 and B78) and 1kb markers (1kb, Promega). A positive radiation hybrid was expected to show a nested PCR product of 863bp (arrowhead). Additionally, a second product, approximately 150bp shorter, was observed consistently. Radiation hybrids 8, 49, 59, 70, 73, 74, 80, 84, 87, 109, 135, 190, 304, 306, 308 and 311 proved positive. The very faint band amplified from hybrid 70 was not reproducible in the duplicate screen. All other positives were confirmed in the second assay (data not shown). Submitting this data to the radiation hybrid mapping webpage placed *sox10* 0cR from z8492 with a high lod score of 17.6.


A
1kb 1 2 4 5 6 8 9 11 13 16 18 19 23 25 27 29 30 33 36 39 40 41 1kb
2.0kb
1.0kb
0.75kb
B
1kb 47 48 49 50 52 54 57 59 63 65 66 70 73 74 79 80 83 84 85 86 87 88 1kb
2.0kb
1.0kb
0.75kb
C
1kb 89 91 92 96 97 98 101 104 105 106 108 109 114 117 119 121 123 125 132 135 136 137 1kb
1.5kb
1.0kb
0.75kb
0.5kb

**D**


1kb 138 150 151 152 153 167 169 174 175 176 178 182 183 184 190 300 301 302 303 304 305 306 1kb
2.0kb
1.0kb
0.75kb

**E**


1kb M
S11-S13 mix AB9 B78
/
308 309 310 311 312
S12-S13 mix AB9 B78
1kb M
1.5kb
1.0kb
0.75kb
0.5kb

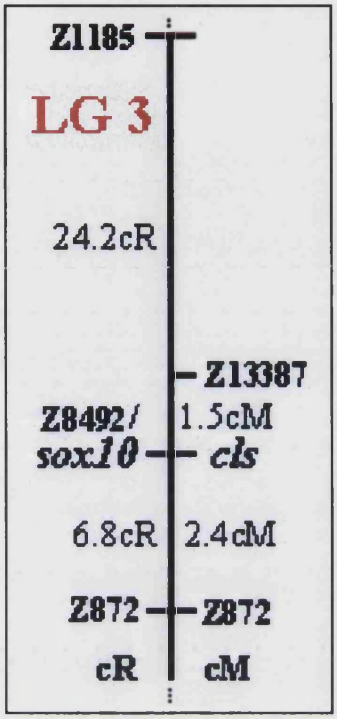



**Figure 5.5:** Map of the distal end of LG 3 containing the *sox10* locus

On the left, three markers mapped on the radiation hybrid panel LN54 are shown and their distance to *sox10* in cR. On the right, microsatellite markers from a linkage map of the equivalent region are shown. The marker Z872 has been placed on both maps and is 2.4cM (~1.4Mb) or 6.8cR (~1Mb) from the *cls* mutant locus or *sox10*, respectively. *Sox10* was shown to be linked to LG 3 with a high lod score of 17.6, 0cR away from Z8492 on the radiation hybrid panel LN54, whereas *cls* was mapped to a 3.9cM region between markers Z13387 and Z872 by linkage mapping of 274 meioses (Dutton et al., 2001).

the radiation hybrid panel are combined in one map. The presence of some markers in both maps close to *cls/sox10* shows that the *cls* mutation is located in the same region as *sox10*, consistent with the proposal that *cls* encodes *sox10*.

## Discussion

We have demonstrated close linkage between the *cls* mutant locus and *sox10*. This result was unambiguous and strongly in support of the hypothesis that *cls* encodes *sox10*.

This method of mapping is based on a PCR screen, which, although very quick, is very sensitive to slight changes in conditions and DNA contamination of the template or solutions involved. For this reason, the screen was carried out in duplicate. Furthermore, variability of the intensity of PCR products between individual radiation hybrid cell lines can be explained by different amounts of radiated zebrafish genomic DNA retained by the recipient mouse cell line.

To ensure specificity and detection of fainter products, a nested PCR was carried out, which consistently amplified two bands. There are two possibilities for the origin of this second band. Although the primer pair S11-S13 had amplified a single PCR product, the nested primer S12 could have misprimed and amplified a slightly shorter product alongside of the one with the expected size in the nested PCR reaction. Alternatively, the reverse primer S13 in the first PCR reaction might have misprimed resulting in two first round PCR products. Both products were again amplified in the nested PCR reaction since they both contained the S13 primer sequence at their 3' ends and S13 was also used as a nested primer. The latter explanation seemed more likely, since there were also two faint products observed in the AB9 control reaction of the first round of PCR amplification (Figure 5.4E). As expected they were both slightly larger than each of the two nested products and were only visible in the reaction with 100% zebrafish genomic DNA, but not in the mix containing 10% zebrafish and 90% mouse genomic DNA. Even though these two bands were obtained for each positive radiation hybrid, they were observed consistently and improved scorability immensely. To prevent the amplification of such a second product in the future, one should set up a fully nested second round PCR reaction. The latter uses two nested primers instead of one nested and one first round primer.

Radiation hybrid 70 amplified a very faint *sox10* specific product in the first nested assay, which was not reproducible in the duplicate screen. However, all positive hybrids consistently produced fainter bands in the second nested assay, most likely due to suboptimal conditions caused by a mixture of Taq DNA polymerases used. Thus, amplification from hybrid 70 might have failed or the amount of product obtained was below the detection sensitivity limit of a standard agarose electrophoresis. A discrepancy of 1 hybrid producing varying results is considered acceptable according to the radiation hybrid webpage (http://zfin.org/ZFIN/, LN54 mapping panel) in which is stated that a marker should be reconsidered, if 5 or more hybrid cell lines produce varying results.

The lod score of 17.6, with which *sox10* was placed 0cR from Z8492, was highly significant. It placed *sox10* with great probability in close proximity to the *cls* mutant locus on LG 3.

Very recently, Tom Carney showed that injection of a 30kb PAC clone containing *sox10* into 1-16 cell embryos was able to rescue the *cls* mutant phenotype (T. Carney, unpubl. data). Thus, the distance between *cls* and *sox10* could in fact be decreased to 30kb.

Although these results do not prove that *cls* encodes *sox10*, it is strongly consistent with the latter and encouraged us to look directly for lesions in *sox10* in different *cls* mutant alleles.

# Chapter 6 – Identification of mutant lesions in *cls* alleles

## *Introduction*

In the previous chapter, we showed tight linkage between *sox10* and the *cls* mutant locus. This was consistent with our hypothesis that *cls* encoded *sox10*. However, it was necessary to further test this hypothesis by trying to identify a mutant lesion in at least one of our *cls* alleles.

As described in the Introduction, the $cls^{t3}$ mutant first arose spontaneously within the *rose* stock kept in Tuebingen. Four additional *cls* alleles, $cls^{ty22f}$, $cls^{te275}$, $cls^{tw2}$ and $cls^{tw11}$, were identified in the Tuebingen mutagenesis screen and two $cls^{m241}$ and $cls^{m618}$, in a similar screen in the Driever lab in the USA (Driever et al., 1996; Kelsh et al., 1996; Malicki et al., 1996). Apart from $cls^{m241}$, which might have been lost, all 6 other alleles are currently available in Bath.

## Methods for identifying mutant lesions

There are at least three approaches to search for a mutant lesion. If a mutant phenotype is caused by a large deletion within the transcribed region, a difference in size of mutant and wild-type mRNA transcripts might be detectable in a Northern Blot.

Similarly, where point mutations or deletions are expected, a sequence difference may be detected in PCR amplified exon sequences of candidate genes by single-strand conformational polymorphism (SSCP) analysis.

Alternatively, point mutations or small deletions can be detected by direct sequencing of mutant and wild-type alleles and subsequent sequence comparison.

Which method is chosen depends on the type of mutation expected. All *colourless* alleles, apart from $cls^{t3}$, were isolated from large-scale mutagenesis screens and contain ENU-induced mutations. The ENU mutagen is known to mostly introduce point mutations rather than large deletions and thus a direct sequencing approach was chosen. This method is technically much easier compared to SSCP and methodology and equipment were readily available in our laboratory.

The 1455bp coding region was too long to be sequenced with one primer from either end. Hence, 4 primer pairs were designed spanning the entire coding region (Figure

6.1). RT-PCR would amplify 4 overlapping fragments between each primer pair from wild-type and mutant alleles. Sequence comparison of corresponding wild-type and mutant allele fragments might then reveal the site of mutation.

This approach bears 3 caveats. Firstly, it is important to avoid contaminating wild-type embryos when collecting homozygous mutant embryos as a source for mRNA. Even one missorted wild-type embryo within mutants would result in amplification of a genotypically mixed sample and might return ambiguous sequence data. Wild-type embryos start to pigment at approximately 25hpf (Kelsh et al., 1996), but at that early stage it is very difficult to sort mutant embryos reliably. On the other hand, mutant embryos should be collected as early as possible, since the *sox10* expression level in *cls* mutants decreases noticeably after 24hpf. Thus, total RNA from different *cls* alleles was isolated from homozygous mutant embryos between 25hpf and 37hpf (prim6-prim25).

Secondly, great care had to be taken to avoid cross contamination when setting up PCR reactions from wild-type and various mutant alleles. Thus, solutions including primers, buffers, magnesium chloride and milliQ-water used in RT-PCR from mutant alleles were prepared freshly and kept separately. Furthermore, all PCR reactions were set up using filter tips to avoid contamination originating from pipettes.

Finally, PCR errors introduced by the Taq polymerase had to be distinguished from actual sites of mutations. One of the two possible approaches is to clone PCR fragments and sequence several clones. In contrast to a true point mutation, it is unlikely that a PCR error would be found in all sequenced clones in the same location. For the same reason, the alternative approach is to directly sequence the entire pool of PCR fragments. Only the aberrant base of the point mutation would be found in all templates and thus, be detected as a strong band on a sequencing gel.

## *Results*

The method chosen for the search of mutant lesions involved RT-PCR of 4 overlapping fragments from wild-type and various mutant alleles and sequence comparison of equivalent wild-type and mutant regions.

Thus, total RNA was prepared from 25-37hpf *cls* mutant embryos using TRI reagent. Numbers of mutant embryos varied between alleles depending on the size and quality of the batch of eggs. RNA pellets were resuspended in a suitable volume of

# Primers to search for mutant lesions

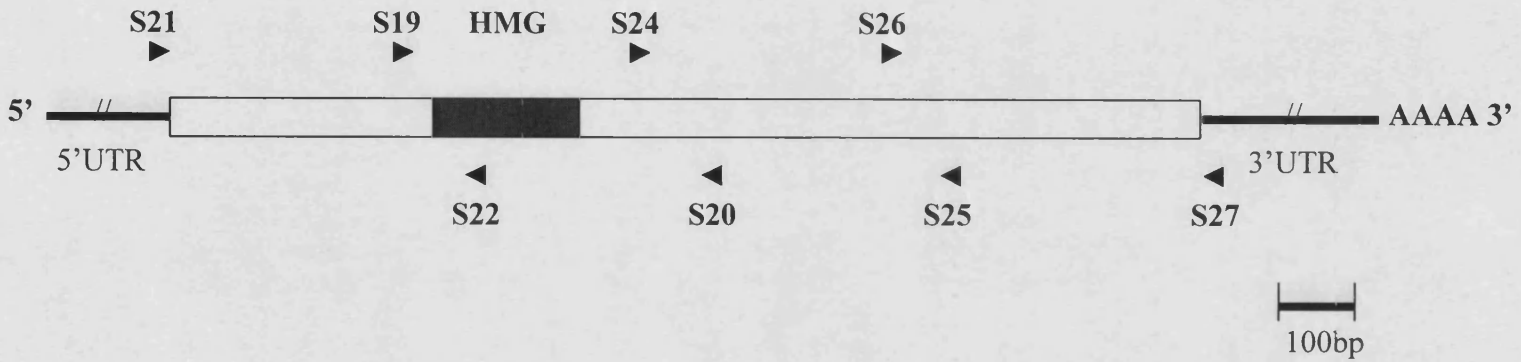



**Figure 6.1:** Location of primers used to sequence wild-type *sox10* and *cls* mutant alleles in search for mutant lesions.

The coding (open box) and untranslated regions (5' UTR, 5' untranslated region; 3' UTR, 3' untranslated region) of *sox10* are shown together with the HMG box (black). Four overlapping primer pairs were designed, which span the entire length of the coding region, S21-S22, S19-S20, S24-S25 and S26-S27. RT-PCR with total RNA from wild-type and homozygous *cls* mutant embryos was performed with each primer pair. Sequences of equivalent wild-type and mutant fragments were compared to identify mutant lesions. The scale bar represents the length of a 100bp fragment.

DEPC-treated water (Appendix 6.1). Agarose gel electrophoresis of samples of total RNAs confirmed the good quality of the RNA preparations (Figure 6.2). In each sample, the 3 ribosomal RNA bands, 23S, 16S and 5S, were observed. A slight background smear indicated the presence of mRNA transcripts. For each allele, approximately 3μg total RNA were reverse transcribed using an oligo(dT) primer. The RT-PCR conditions and amplification program for each of the 4 primer pairs were determined in advance by using the PCR optimisation scheme (Materials and Methods). This scheme tested different combinations of magnesium, dNTP and primer concentrations on cDNA derived from 19hpf wild-type total RNA. The conditions, which produced the strongest band with the least non-specific by-products were chosen for mutant RT-PCRs (Figure 6.3). For each primer pair, RT-PCR reactions were set up with the optimised conditions and cDNA from each of the mutant alleles, *cls*$^{tw1}$, *cls*$^{tw2}$/1, *cls*$^{t3}$/1, *cls*$^{tw11}$, *cls*$^{ty22f}$ and *cls*$^{m618}$/1, and a "no cDNA" negative control (Materials and Methods). Expected products ranged from 432bp to 547bp depending on the primer pair (Figure 6.4). No such bands were observed in the "no cDNA" controls. Equivalent fragments derived from different mutant alleles were always the same size and thus, were unlikely to contain larger insertions or deletions. All 4 primer pairs amplified products from *cls*$^{tw2}$, *cls*$^{t3}$, *cls*$^{tw11}$ and *cls*$^{m618}$. Very faint products were observed from *cls*$^{ty22f}$ with primer pair S19/S20 and S21/S22, but no products were detectable with primer pair S24/S25 and S26/S27. The opposite was true for *cls*$^{tw1}$; no bands were obtained with primers S19/S20 and S21/S22 and faint bands with S24/S25 and S26/S27 (Figure 6.4).

Mutant PCR products and the equivalent wild-type fragments from the PCR optimisation were purified using the QIAquick PCR Purification Kit (Qiagen). After the DNA was bound to the column, an additional wash with 35% guanidine hydrochloride was included. This wash was not part of the standard protocol, but was suggested in the instruction manual to completely remove primer dimers longer than 20bp. This step was important, since these purified PCR products were subsequently included in a sequencing reaction, in which either forward or reverse PCR primers were used as a sequencing primer. Presence of primer dimers might cause failure of the sequencing reaction by binding the sequencing primer. The concentration of purified samples was estimated by agarose gel electrophoresis and comparison to a quantitative DNA marker (data not shown). Suitable amounts of each product were sent off for automated sequencing with the appropriate forward and reverse PCR primer.

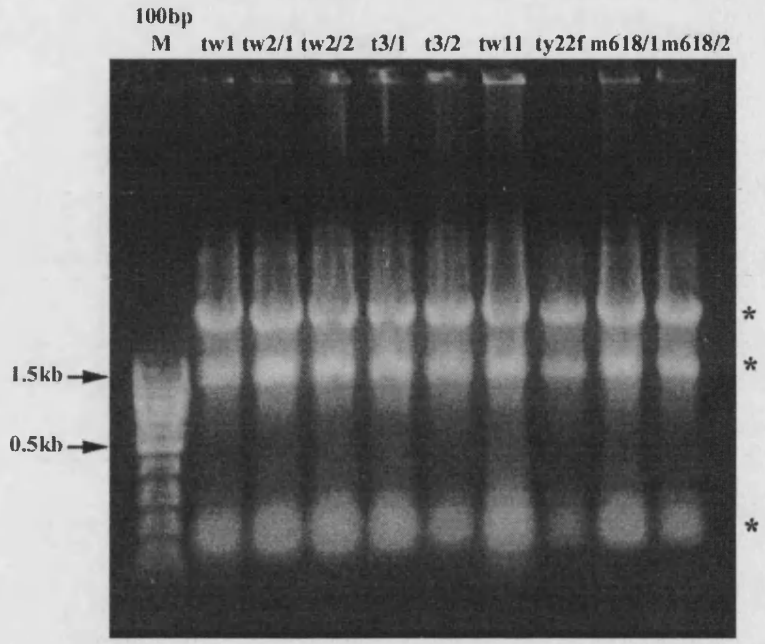



**Figure 6.2:** Agarose gel electrophoresis of total RNA preparations from *cls* mutant alleles

The quality of total RNA from 25-37hpf homozygous *cls* mutant embryos was judged to be good by the presence of 3 distinct ribosomal RNA bands, 23S, 16S and 5S (asterisks), in combination with a slight background smear corresponding to mRNA transcripts. RNA from the following mutant alleles was analysed: $cls^{tw1}$ (tw1), 2 separate batches of $cls^{tw2}$ (tw2/1 and tw2/2), 2 separate batches of $cls^{t3}$ (t3/1 and t3/2), $cls^{tw11}$ (tw11), $cls^{ty22f}$ (ty22f), 2 separate batches of $cls^{m618}$ (m6182/1 and m618/2). A 100bp marker (100bp M, Promega) was used to estimate the size of the bands.

**Figure 6.3:** PCR optimisations for primer pairs S19-S20, S21-S22, S24-S25 and S26-S27

(A-D) Agarose gel electrophoresis of PCR reactions with wild-type RNA and primer pairs S19-S20 (A), S21-S22 (B), S24-S25 (C) and S26-S27 (D) are shown to test optimum PCR conditions. Expected PCR products are shown (arrowheads in A-D) for each primer pair together with 100bp markers (100bp M, Promega; 100bp M1, Promega; 100bp M2, Generuler). Lanes 1-12 (A-D) correspond to 12 different combinations of magnesium, dNTP and primer concentrations (reactions "111"-"722" in PCR optimisation scheme, Materials and Methods). For each primer pair, the conditions which produced the strongest product with the least non-specific by-products were used to set up RT-PCR reactions with RNA from homozygous *cls* embryos. For primer pairs S19-S20 (A) and S21-S22 (B) reaction condition 4 (white asterisk) was chosen (reaction "122" in the PCR optimisation chart), whereas for primer pairs S24-S25 (C) and S26-S27 (D) reaction condition 3 (white asterisk, reaction "121").

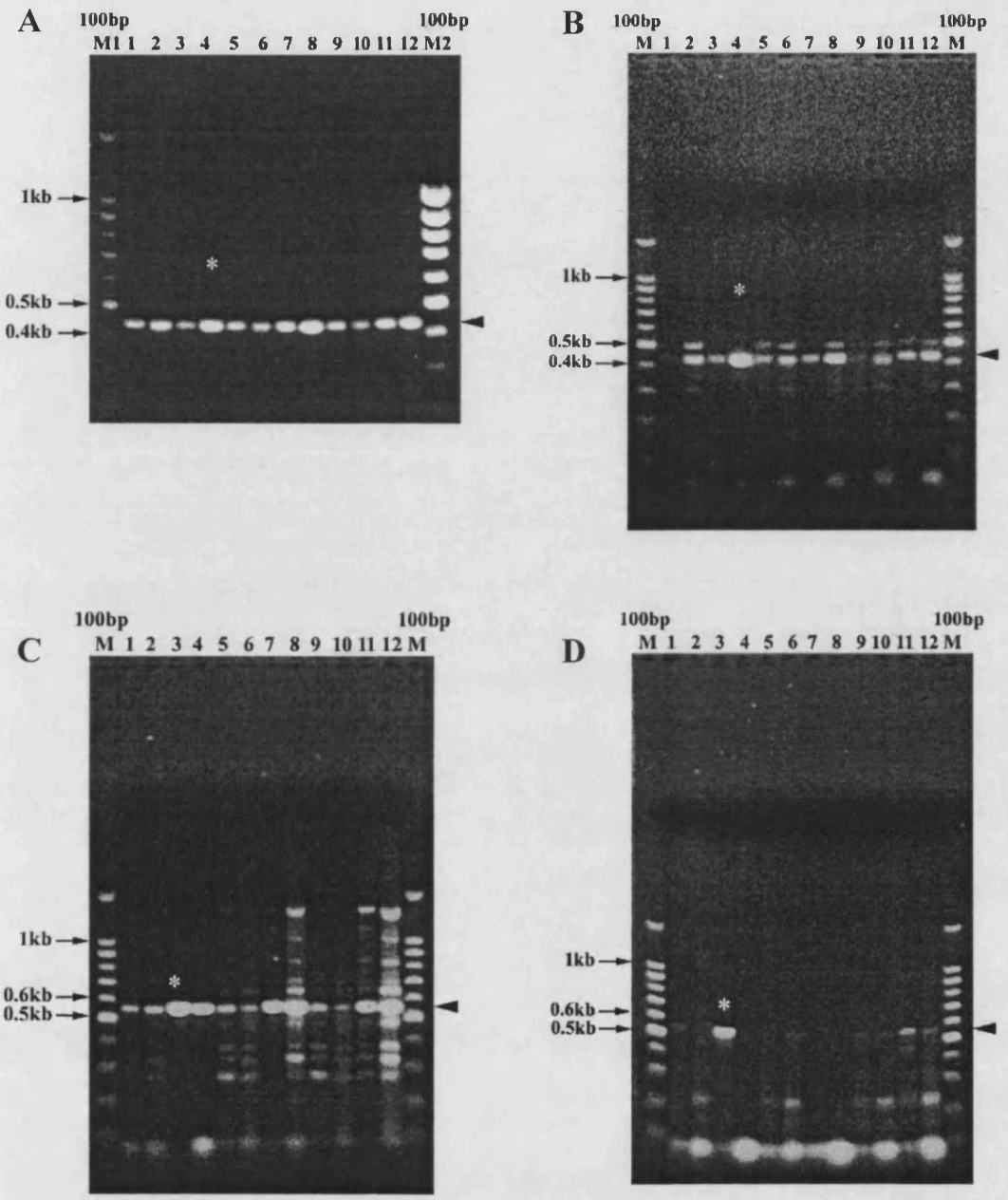
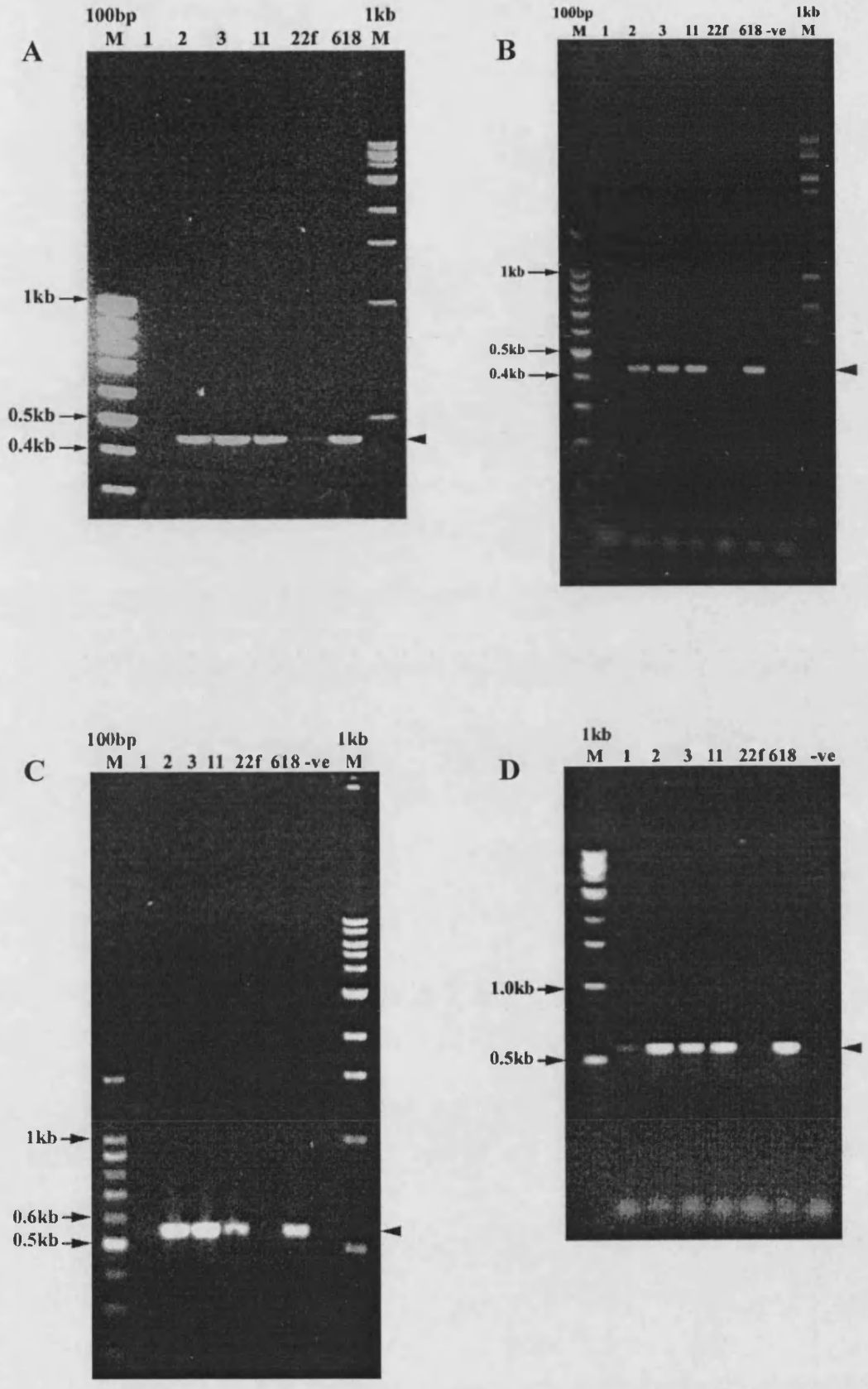

A
100bp
100bp
M1 1 2 3 4 5 6 7 8 9 10 11 12 M2
1kb
0.5kb
0.4kb
B
100bp
100bp
M 1 2 3 4 5 6 7 8 9 10 11 12 M
1kb
0.5kb
0.4kb
C
100bp
100bp
M 1 2 3 4 5 6 7 8 9 10 11 12 M
1kb
0.6kb
0.5kb
D
100bp
100bp
M 1 2 3 4 5 6 7 8 9 10 11 12 M
1kb
0.6kb
0.5kb

**Figure 6.4:** RT-PCR from *cls* mutant alleles

(A-D) Samples of RT-PCR reactions with *cls* mutant alleles (lane 1, $cls^{tw1}$; lane 2, $cls^{tw2}$; lane 3, $cls^{t3}$; lane 11, $cls^{tw11}$; lane 22f, $cls^{ty22f}$; lane 618, $cls^{m618}$) and 4 primer pairs are shown together with 100bp markers (100bp M, Generuler, A; Promega, B and C) and 1kb markers (1kb M, NEB). (A) Amplification between primers S19 and S20 obtained a band of the expected size of 435bp (arrowhead) from mutant alleles $cls^{tw2}$, $cls^{t3}$, $cls^{tw11}$ and $cls^{m618}$. Only a faint product was observed from $cls^{ty22f}$, and no band was detectable from $cls^{tw1}$ or the "noDNA" negative control (not shown). (B) Primer pair S21 and S22 amplified a 432bp product (arrowhead) from mutant alleles $cls^{tw2}$, $cls^{t3}$, $cls^{tw11}$ and $cls^{m618}$, but again, only a barely visible band from $cls^{ty22f}$ and no band was detectable in the negative control (lane "–ve") or from $cls^{tw1}$. (C) shows 547bp PCR products (arrowhead) derived from the primer pair S24 and S25 in an RT-PCR reaction from $cls^{tw2}$, $cls^{t3}$, $cls^{tw11}$ and $cls^{m618}$. Here, $cls^{tw1}$ produced a very faint band, while no band was generated from $cls^{ty22f}$. No band was seen in the negative control (lane "-ve") as expected. (D) RT-PCR reactions with primer pair S26 and S27 obtained a 546bp product from the same alleles as primer pairs S24 and S25 (C).


A
100bp
1kb
M 1 2 3 11 22f 618 M
1kb
0.5kb
0.4kb
B
100bp
1kb
M 1 2 3 11 22f 618 -ve M
1kb
0.5kb
0.4kb
C
100bp
1kb
M 1 2 3 11 22f 618 -ve M
1kb
0.6kb
0.5kb
D
1kb
M 1 2 3 11 22f 618 -ve
1.0kb
0.5kb

Sequences of equivalent fragments of wild-type and each mutant allele were aligned using Pileup (gcg package). The $cls^{tw1}$ sequence from the faint products obtained with S24/S25 and S26/S27, and the $cls^{ty22f}$ sequence from primer pair S19/S20 and S21/S22 matched the wild-type sequences completely. The same was found to be true for all 4 overlapping $cls^{t3}$ fragments. Thus, we did not identify any mutant lesion in these regions of the $cls^{tw1}$, $cls^{t3}$ and $cls^{ty22f}$ alleles.

In contrast, putative point mutations were identified for $cls^{m618}$, $cls^{tw2}$ and $cls^{tw11}$. Thus, the $cls^{m618}$ sequence showed one non-silent single base pair mismatch from WT, which was identified in the HMG box fragment (primer pair S19-S20). Even after sequencing both strands multiple times, a thymine (T) at position 800 was consistently replaced by an adenine (A) resulting in a leucine to glutamine (L142Q) change (Figure 6.5). To confirm this result, the second batch of $cls^{m618}$ total RNA (m618/2, Figure 6.2) was reverse transcribed and the RT-PCR product between primers S19 and S20 sequenced. Again, the same mutant lesion was identified.

Similarly, a point mutation was detected in the overlap between fragments S24-S25 and S26-S27 in $cls^{tw2}$ and $cls^{tw11}$. In both alleles, an A-to-T transversion at position 1501 resulted in an in-frame stop codon (TAA) instead of a lysine (K376X, Figure 6.5). This prematurely truncates the protein just upstream of the putative transactivation domain. Again, the lesion was confirmed by multiple sequencing runs of the S24-S25 and S26-S27 PCR products of both alleles and also by sequencing the equivalent RT-PCR product from the second $cls^{tw2}$ batch (tw2/2, Figure 6.2).

## Discussion

We have described the successful identification of mutant lesions in 3 *cls* alleles. As expected, all 3 mutations induced by the chemical mutagen ENU were point mutations.

$cls^{m618}$ was found to be caused by a point mutation within the HMG domain. This replaced a leucine with a glutamine at residue 142, located in the second of 3 α-helices that give the HMG domain its L-shaped structure. The exact function of this residue has not been elucidated. However, it is fully conserved across all members of the *Sox* gene family and even in the outgroups LEF1 and TCF1 from mouse and C. elegans (Bowles et al., 2000), which might indicate an important structural or functional role. Furthermore, the same residue in human SRY (Leu-39) was shown to be important

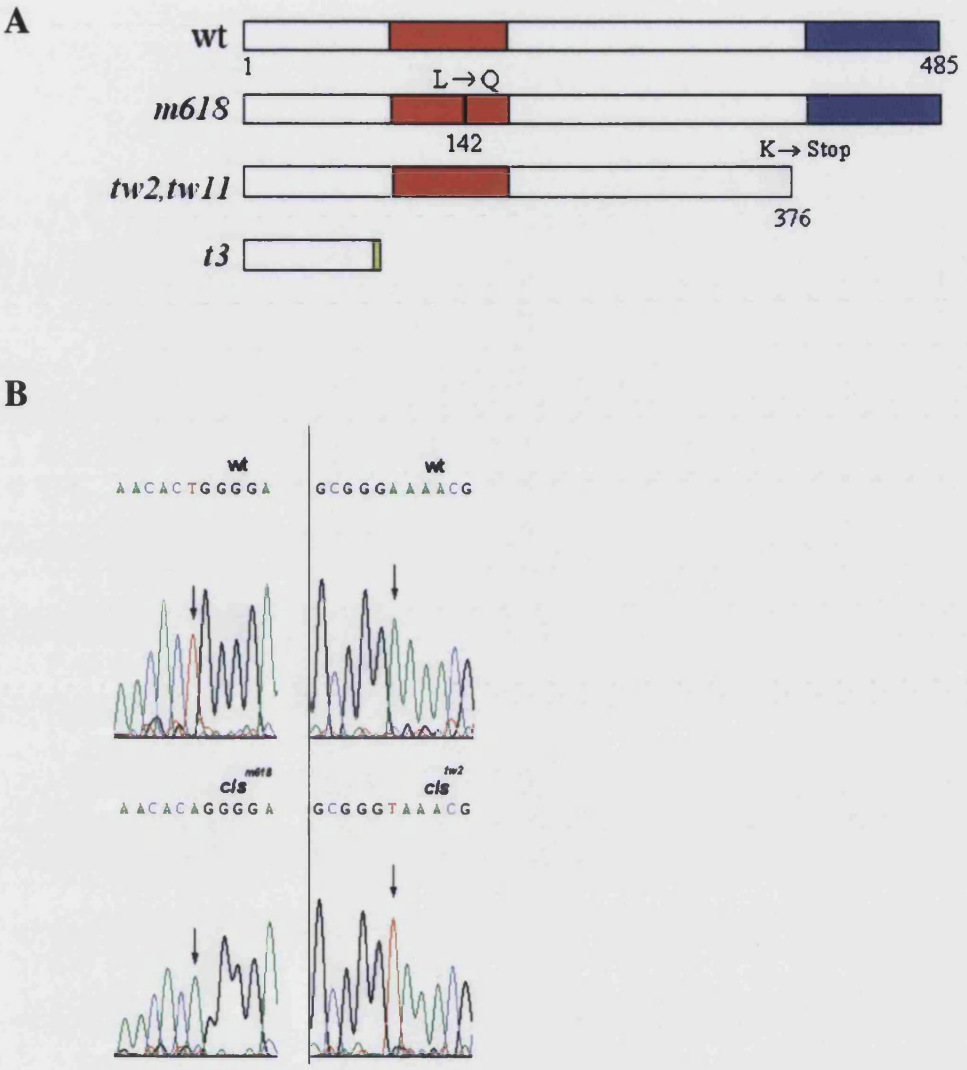



**Figure 6.5:** The mutant lesions in $cls^{m618}$, $cls^{tw2}/cls^{tw11}$ and $cls^{t3}$

(A) The schematic compares wild-type (wt, residues 1-485) to mutant (*m618*, *tw2*, *tw11*, *t3*) Sox10 proteins. In *m618*, a T425A transversion within the HMG domain (red) replaced a leucine (L142) with a glutamine (Q). In *tw2* and *tw11*, an A1126T transversion introduced a Stop codon truncating the protein just N-terminal to the transactivation domain (blue). *t3* was found to be disrupted by a 1.4kb transposon (T. Carney; Dutton, Pauliny et al. 2001). After introducing a C-terminal extension of 8 novel amino acids (yellow), the protein is truncated just N-terminal to the HMG domain. (B) Chromatogram traces show the nucleotide changes (arrows) causing the *m618* and *tw2/tw11* mutant alleles.

in the maintenance of orientation of the long and short arms of the L-shaped HMG domain structure (Werner et al., 1995). The mutation is thus likely to disrupt the three-dimensional structure of the DNA binding domain reducing the ability of Sox10 to bind to its DNA target sites.

In collaboration with Bill Pavan, National Institute of Health, Bethesda, MD, the transcriptional abilities of wild-type and mutant zebrafish Sox10 proteins were tested in luciferase reporter assays (M. Brady, unpubl. data). In transiently transfected HeLa cells, the ability of Sox10 to induce the *nacre* promoter was measured by the activation of a luciferase reporter gene regulated by the *nacre* promoter as previously described in Potterf et al. (2000). The transcriptional activation ability of Sox10 is measured as the fold increase relative to a promoterless basic reporter plasmid. Wild-type Sox10 produced a 46.14 fold induction of luciferase activity, whereas L142Q only a 0.83 fold induction. The loss of function demonstrated in this *in vitro* system confirms this residue's importance for Sox10 as a transcription factor observed in the mutant phenotype. If the mutation renders Sox10 unable to bind to its consensus sites in the promoter regions of downstream targets, this might also abolish its synergistic abilities. It has been demonstrated that both the synergy and the HMG domain are necessary for Sox10 to act cooperatively with partner transcription factors like Pax3 or Tst-1/Oct6/SCIP (Kuhlbrodt et al., 1998a; Bondurand et al., 2000).

Alleles $cls^{tw2}$ and $cls^{tw11}$ contained the same mutant lesion in the exact same location. At first, this seemed a surprising result, but it could be explained by the fact that they were derived from the same mutagenised family founder male, W. Thus, very likely they represented two independent isolations of the same mutation. The point mutation introduced a stop codon, 18 residues upstream of the putative transactivation domain, which resulted in a protein truncated by 109 residues. Due to the loss of the transactivation domain, the protein's function is expected to be severely affected. The *in vitro* analysis of the K376X mutation (W. Pavan, unpubl. data) showed only a 0.87 fold induction of the luciferase reporter construct in comparison to 46.14 fold induction by the wild-type protein. This drastic decrease in transcriptional activity was comparable to the L142Q mutation (0.83 fold induction).

Thus far, we were unable to identify a mutant lesion in the $cls^{ty22f}$ coding region. Faint PCR products could only be obtained from the first two overlapping fragments. This observation might be explained in several ways. Firstly, only a small number of embryos were available from which total RNA was isolated (Table 6.1). Although it

was attempted to compensate for this by resuspending the RNA pellet in a smaller volume and using a larger volume of cDNA to set up PCR reactions, it might have caused fainter products. Secondly, one of the primer sites might be changed by the mutant lesion causing the amplification between primer pair S24-S25 or S26-S27 to fail. Even so, this could not explain why PCR reactions with both primer pairs failed to produce products. Alternatively and most likely, it is possible that the products observed were derived from a small contamination of wild-type RNA. This might have been caused by a missorted wild-type embryo within the *cls* siblings or whilst setting up the PCR reaction. Consistent with this possibility, the sequence of the faint products matched wild-type *sox10* completely. Furthermore, *sox10 in situ* hybridisation on $cls^{ty22f}$ revealed overall reduced expression levels in a quarter of the embryos, presumably corresponding to homozygous siblings (Figure 6.6). Taken together, these observations indicate that $cls^{ty22f}$ is caused by a further mutant lesion different to L142Q and K376X. It might be located outside the coding region in one of the regulatory elements, in which case one approach to find the mutant lesion is to compare the promoter regions of wild-type and $cls^{ty22f}$ mutants by sequencing. Alternatively, the transcript might be destabilised due to the lesion affecting intron/exon boundaries and thus correct splicing of the transcript. To identify such a lesion, one could sequence across intron/exon boundaries in genomic DNA of $cls^{ty22f}$ mutants.

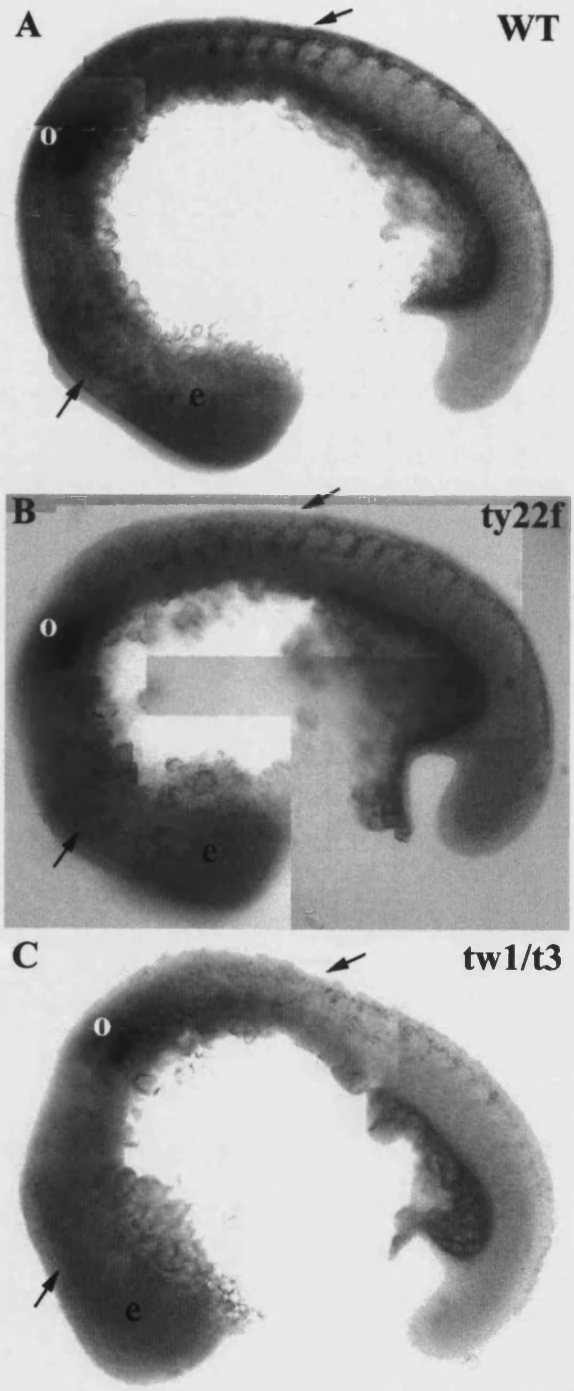
The search for a mutant lesion by sequencing the 2 faint products obtained from $cls^{tw1}$ was unsuccessful. There are a two explanations why only the two C-terminal fragments, which matched wild-type *sox10* completely, were amplified. Either, mutant transcripts were produced, but the PCR reactions with the N-terminal primer pairs S19-S20 and S21-S22 failed altogether and hence, no products were obtained or PCR products might have been present, but too faint to be detectable by agarose gel electrophoresis. Alternatively, the mutant lesion could be located in a regulatory element and hence no transcripts are synthesised or only at a drastically reduced level. Thus, the PCR products obtained might have resulted from wild-type contaminations. Consistent with this hypothesis, *sox10 in situ* hybridisation on $cls^{tw1}$ embryos showed that in a quarter of embryos at any stage, expression was barely visible apart from the expression in the otic vesicle (Figure 6.6). This is different to most other alleles, in which *cls* mutants are indistinguishable from wild-type siblings up to 24hpf.
Only recently, we discovered that $cls^{tw1}$ and $cls^{t3}$ were identical. For reasons unknown, $cls^{t3}$ had been renamed $cls^{tw1}$ in the Tuebingen stock centre. Furthermore, T. Carney

**D**

| | $cls^{ty22f}$/1 r/t (%) | $cls^{ty22f}$/2 r/t (%) | $cls^{tw1}$/1 r/t (%) | $cls^{tw1}$/2 r/t (%) |
|---|---|---|---|---|
| **2 somites** | 9/36 (25%) | nd | 10/27 (37%) | 8/32 (25%) |
| **18 somites** | 10/38 (26%) | 7/25 (28%) | 8/30 (27%) | 7/26 (27%) |
| **24 hpf** | 3/12 (25%) | 3/21 (14%) | 3/20 (15%) | 6/25 (24%) |

**Figure 6.6:** The $cls^{ty22f}$ and $cls^{tw1}$ ($cls^{t3}$) mutant alleles

(A-C) *In situ* hybridisation with *sox10* on 18 somite stage embryos derived from a heterozygous $cls^{ty22f}$ cross shows a reduction (B, black arrows) of *sox10* transcripts in approximately a quarter of embryos, presumably homozygous mutants, whereas equivalent $cls^{tw1}$ ($cls^{t3}$) embryos reveal an even more severe reduction (C, black arrows) of *sox10* staining. In comparison, *cls* mutants of other alleles at this stage are still indistinguishable from wild-types (A). The expression in the otic vesicle (o) in $cls^{ty22f}$ (B) is comparable to wild-types (A) and significantly reduced in $cls^{tw1}$ (C). (D) The number of embryos with reduced *sox10* expression (r) from 2 independent hybridisation reactions ($cls^{ty22f}$/1, $cls^{ty22f}$/2 and $cls^{tw1}$/1, $cls^{tw1}$/2) was compared to the total number (t) of embryos included in each staining reaction and the percentage (%) calculated for different stages (2 somite, 18 somite and 24hpf stage). If the siblings with reduced *sox10* staining corresponded to homozygous *cls* mutants, a percentage value of approximately 25% of the total number of embryos is expected in all cases. e, eye; hpf, hours post fertilisation; nd, not determined; o, otic vesicle; r, reduced; t, total.


A
WT
o
e
B
ty22f
o
e
C
tw1/t3
o
e

recently showed by sequencing genomic DNA from *cls*$^{t3}$ embryos, that *cls*$^{t3}$ was caused by a 1.4kb insertion of a transposable element at nucleotide 255 of the transcript, a position 20 amino acids upstream of the HMG domain (Genbank accession number AF404490; Dutton et al., 2001b). The *cls*$^{t3}$ lesion results in premature truncation of the N-terminal domain with an extension of 8 novel amino acids. The protein thus only consists of what has been defined as the synergy domain (residues 1-89; Kuhlbrodt et al., 1998a). Since products of expected sizes were amplified between all 4 primer pairs from *cls*$^{t3}$ cDNA, at least the band between primer S21 and S22 must have been contamination from either wild-type or other mutant cDNAs. With the knowledge of *cls*$^{tw1}$ and *cls*$^{t3}$ being identical, it is more likely that all bands amplified resulted from wild-type DNA contamination, which helps to explain our confusing and inconsistent results obtained from these two mutant PCR reactions. It will be interesting to characterise this allele in greater detail *in vivo* (rescue experiments) and *in vitro* (luciferase reporter assays). However, it is likely to be a null, since the DNA binding domain, which also contains two nuclear localisation signals (Suedbeck and Scherer, 1997), and the transactivation domain are missing.

Preliminary phenotypic comparisons by project students did not reveal greatly significant differences between these alleles. The effect of *cls* on the pigment phenotype was assessed by counting melanoblasts in 22-24hpf wild-type and *cls* mutants labelled by *dopachrome tautomerase* (*dct*) *in situ* hybridisations (Sanders, 2000). This marker is expressed in melanoblasts approximately 5 hours prior to melanisation (Kelsh et al., 2000c). In a similar project, the effect of *cls* on the peripheral nervous system was investigated. The reduction of DRG and enteric neurons in 3dpf and 5dpf *cls* mutants was assessed by counting neurons labelled with the pan-neuronal anti-Hu antibody (Ford, 2000). Some allelic differences that students observed in various cell-types and stages might have been artefacts due to small sample numbers. This suggestion is based primarily on the discrepancies in allelic strengths reported for the *cls*$^{tw1}$ and *cls*$^{t3}$ alleles, which are now known to contain identical mutant lesions. However, overall the severity of phenotypes in alleles investigated was largely comparable and approximately equivalent.

In conclusion, we demonstrated mutant lesions in 3 of the *colourless* alleles and a fourth was identified by T. Carney. Together with the expression studies and the mapping result, this presents a very strong indication that *cls* is *sox10*. Preliminary data from two undergraduate project students showed that all alleles were phenotypically

very similar. The $cls^{t3}$ allele seems highly likely to be a null due to the location of the t3 lesion. However, the severity of effects caused by the $cls^{m618}$ and $cls^{tw2}$ lesions is not quite so clear cut. In the next chapter we attempt to test whether these mutant proteins have any residual function *in vivo*.

# Chapter 7 – Rescue of the *cls* phenotype by heatshock construct injections

## *Introduction*

Can *sox10* rescue the *colourless* phenotype? This is the last of a series of tests to investigate whether *cls* encodes *sox10*. In previous chapters, we have demonstrated that *sox10* and the *cls* mutant locus are very tightly linked, that *sox10* is expressed in a pattern consistent with the mutant phenotype and identified genetic lesions expected to disrupt the Sox10 protein. If *cls* encodes *sox10*, we would also expect to rescue at least partially the *cls* phenotype by expressing wild-type *sox10* in *cls* mutant embryos.

There are generally two methods to ectopically express a gene, either by directly injecting mRNA or a construct containing the cDNA of interest under the regulation of a suitable promoter. If mRNA is injected, the timing of expression of the protein cannot be regulated. It has been previously suggested that RNA injections lead to sufficient RNA and thus protein levels up to midsegmentation stages, while DNA injections become necessary to affect processes that occur after the 20hpf stage (Hammerschmidt et al., 1999). Judged by the first visible defect in *cls* embryos at 24hpf, Sox10 function is likely to be required at a time when ectopic Sox10 protein is unlikely still to be present. Consistent with this, injection of *sox10* mRNA into 1 cell stage embryos of a heterozygous $cls^{tw11}$ cross failed to rescue *cls* mutants (S. Elworthy, unpubl. data).

In contrast, driving ectopic expression of the gene of interest by a known promoter allows controlled expression of the protein at a particular time and duration during development. A heatshock promoter was chosen, since initiation of expression is easily achieved by incubating embryos containing the heatshock construct at the permissive temperature. The heatshock construct is made by cloning the *sox10* coding region into a suitable expression vector, downstream of a heatshock promoter. This construct is then injected into the yolk of early cleavage stage embryos (1-4-cell stage) where it becomes distributed by yolk streaming. Upon raising the incubation temperature of injected fish embryos from 28.5°C to 37°C, transcription of *sox10* from the heatshock promoter is initiated in those cells that have received the construct.
To achieve rescue, it is important that Sox10 is expressed at the correct concentration and time required. To maximise the degree of rescue, we thus anticipated having to

optimise the amount of injected construct per embryo, and the time and number of heatshock treatments. We chose to score rescue based on the rescue of melanophores. This neural crest derivative is severely affected in *cls* embryos of all strong alleles, in which pigmented melanophores with stellate wild-type morphology are never observed. In wild-type embryos, melanophores are easily detected under the dissecting microscope from approximately 30hpf. Thus, rescue was defined as presence of at least one fully differentiated melanophore with wild-type morphology.

We wanted to establish the severity of a mutant allele in relation to the location of the mutant lesion and asked the question whether the resulting mutant protein retained any function *in vivo*. Results from corresponding *in vitro* studies, discussed in the previous chapter, indicated that $cls^{m618}$ and $cls^{tw2/tw11}$ had no transcriptional activity. We used rescue experiments to characterise these mutant alleles and specifically we tested the prediction that they would fail to rescue the *cls* mutant phenotype.

Finally, we wanted to investigate the possibility of functional redundancy between *sox9* and *sox10*. Although there is no previous evidence for redundancy of Sox proteins, it is not unreasonable to assume that highly conserved proteins might function redundantly when co-localised in the same tissues.
*sox9* and *sox10* are the most closely related members of the *Sox* gene family. Zebrafish *sox9b* and maybe also *sox9a*, were shown to be expressed in the neural crest (E. Chiang, pers. commun.; this study) consistent with reports of mouse *Sox9*, which labels cephalic crest (Ng et al., 1997). To test whether either *sox9a* or *sox9b* could rescue aspects of the *cls* phenotype, we replaced the *sox10* coding region in the expression vector with that of *sox9a* or *sox9b*. The same method and scoring criteria were to be applied as had previously been established for *sox10*.

## *Results*

### Preparation of *sox9* and *sox10* heatshock constructs

To generate heatshock constructs, we chose a strategy of PCR amplification of the Sox protein coding regions and cloning them into a heatshock expression plasmid. Clones were sequenced to select against PCR errors which might have been introduced.
This approach required the design of PCR primers outside the coding region and a suitable vector containing a heatshock promoter. The plasmid pCSHSP was chosen as

the expression vector, since it had been successfully used in rescue experiments with the zebrafish *mitf* homologue, *nacre* (*nac*; Lister et al., 1999). It is based on the pCS2+ plasmid, in which the CMV (cytomegalovirus) promoter had been replaced by a heatshock promoter (HSP; Halloran et al., 2000). This modified vector was kindly provided by D. Raible (Figure 7.1). *Cla* I and *Xba* I restriction sites within the multiple cloning site of the vector were selected for subcloning the *sox10*, *sox9a* and *sox9b* coding regions, since neither enzyme was predicted to cut within these fragments. Thus, primers S21 and S27, 9A-1 and 9A-2 and 9B-1 and 9B-2, located just outside the *sox10*, *sox9a* and *sox9b* coding regions, respectively, were redesigned to include *Cla* I and *Xba* I sites at their 5' ends (Figure 7.2).

RT-PCR amplifications were performed on total RNA derived from 38hpf wild-types and $cls^{m618}$ and $cls^{tw2}$ mutants and primers *Cla* I-S21 and S27-*Xba* I (Materials and Methods; Figure 7.3). The *sox9a* and *sox9b* coding regions were amplified from plasmid minipreparations containing the appropriate full-length cDNAs using primers 9A-1 and 9A-2 or 9B-1 and 9B-2, respectively (Materials and Methods; Figure 7.3). For all PCR reactions, either Herculase or TaqPlus Precision PCR system (Stratagene), both high fidelity DNA polymerases, were used to reduce the frequency of PCR errors. PCR reactions usually produced a single product of the expected size, which was directly subjected to a restriction digestion with *Cla* I and *Xba* I (Promega) in multicore buffer (Promega). However, with the $cls^{m618}$ cDNA, primers *Cla* I-S21 and S27-*Xba* I amplified a non-specific product of approximately 400bp alongside the specific 1614bp band. Thus, the specific product was excised and gel purified using the QIAquick Gel Extraction Kit (Qiagen) before restriction digestion with *Cla* I and *Xba* I (data not shown). Digested coding regions were purified using the QIAquick PCR Purification Kit (Qiagen) and the concentration of eluted samples was estimated by agarose gel electrophoresis and comparison to a quantitative DNA marker (NEB; data not shown). In parallel, the pCSHSP plasmid had also been restriction digested with *Cla* I and *Xba* I and purified by standard phenol/chloroform extraction and ethanol precipitation. Heatshock constructs were created by ligating each coding region into linearised pCSHSP plasmids and transforming the constructs into Gold Super-competent Epicurian Coli cells (Stratagene). A suitable number of white colonies were picked and analysed by restriction digestion for inserts of the expected size (Figure 7.4). Only 4/18 (22%) white colonies from the cloning of wild-type *sox10* contained inserts of the correct size. All 4 clones contained a non-conservative PCR error altering the protein

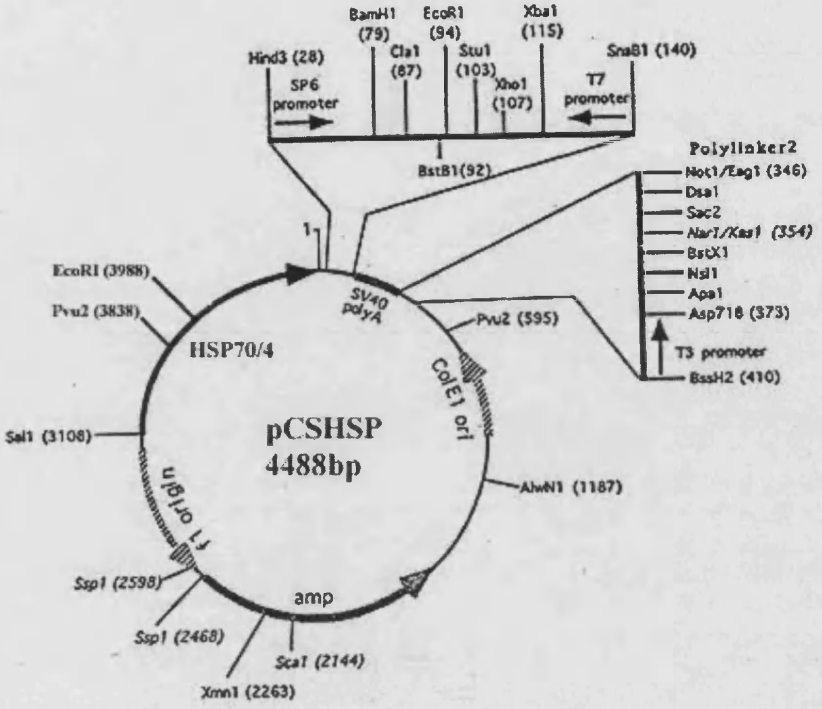



**Figure 7.1:** Plasmid map of pCSHSP

This map was adapted from http://faculty.washington.edu/rtmoon/

This expression vector is based on the pCS2+ plasmid, in which the cytomegalovirus (CMV) promoter had been replaced by a heatshock promoter, HSP70/4 (Halloran et al., 2000). To create the heatshock constructs, the *sox10* coding region was cloned into the *Cla* I (87) and *Xba* I (115) sites within the multiple cloning site, downstream of the HSP. The ampicillin resistance gene (amp) is used as the selectable marker. For RNA injections, the plasmid is linearised with *Asp* 718 (373) and transcribed with SP6 RNA polymerase.

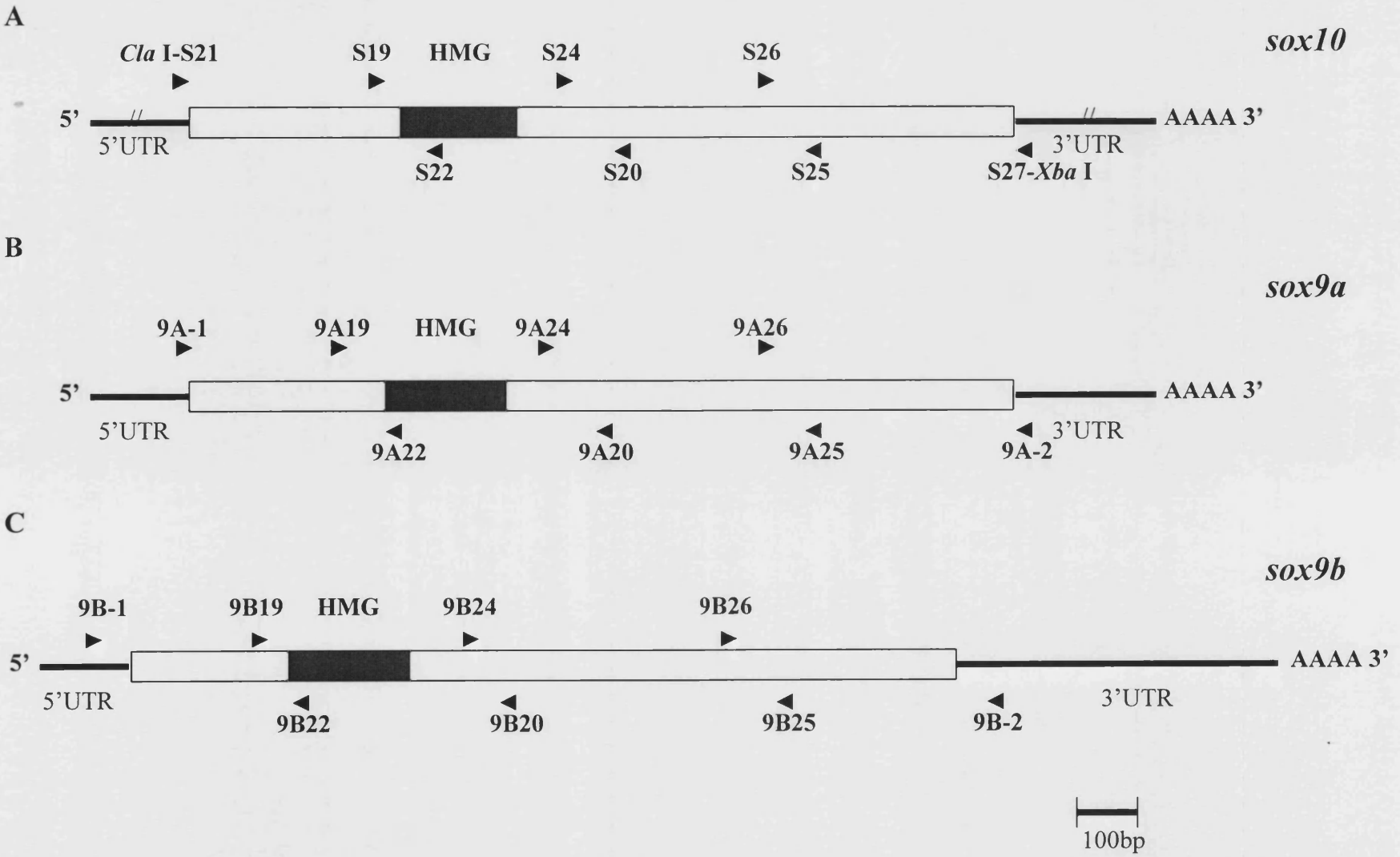
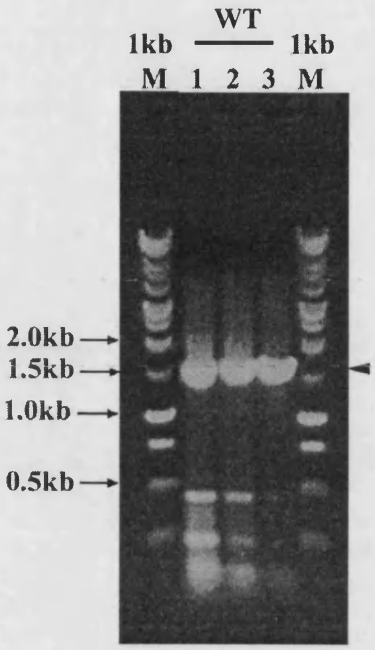
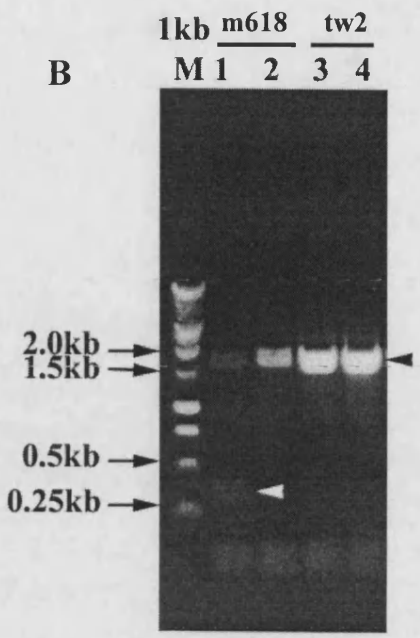

A
sox10
Cla I-S21
S19
HMG
S24
S26
5'
AAAA 3'
5'UTR
3'UTR
S22
S20
S25
S27-Xba I
B
sox9a
9A-1
9A19
HMG
9A24
9A26
5'
AAAA 3'
5'UTR
3'UTR
9A22
9A20
9A25
9A-2
C
sox9b
9B-1
9B19
HMG
9B24
9B26
5'
AAAA 3'
5'UTR
3'UTR
9B22
9B20
9B25
9B-2
100bp

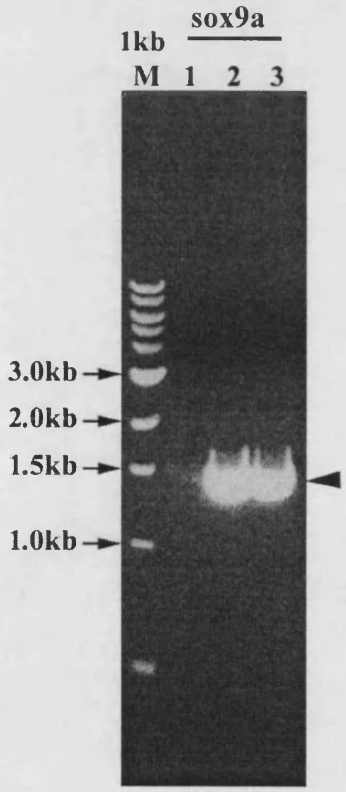
**Figure 7.2:** Primers used to create the *sox10*, *sox9a* and *sox9b* heatshock constructs

(A-C) The coding (open box) and untranslated regions (5' UTR, 5' untranslated region; 3' UTR, 3' untranslated region) of *sox10* (A), *sox9a* (B) and *sox9b* (C) are shown together with their HMG box (HMG, black box). (A) Primer pair *Cla* I-S21 and S27-*Xba* I are located just outside the stop and start codons and thus amplify the entire *sox10* coding region or *cls* mutant allele coding regions if mutant cDNAs were used as templates. After cloning these fragments into pCSHSP vector, the constructs were analysed for possible PCR errors by sequencing with a combination of primers as indicated. (B, C) The *sox9a* (B) and *sox9b* (C) coding regions were amplified from plasmid DNA minipreparations containing the appropriate full length cDNA between primers 9A-1 and 9A-2 or 9B-1 and 9B-2, respectively. These primers had been designed to include *Cla* I (9A-1 and 9B-1) and *Xba* I (9A-2 and 9B-2) restriction enzyme recognition sites at their 5' ends facilitating subsequent cloning into the heatshock plasmid pCSHSP. The sequence of the cloned *sox9a* and *sox9b* heatshock constructs were checked for PCR errors using the sequencing primers as indicated. The scale represents the length of a 100bp fragment.

**Figure 7.3:** PCR of wild-type and mutant *sox10,* and wild-type *sox9a* and *sox9b*
(A, B) To amplify the *sox10* coding region, RT-PCR was performed on total RNA from 38hpf wild-types (A) or 25-37hpf $cls^{m618}$ (B, lane 1 and 2) and $cls^{tw2}$ mutants (B, lane 3 and 4) between primers *Cla*I-S21 and S27-*Xba*I. (A) Reactions were set up with 1μM (lane 1), 0.5μM (lane 2) and 0.2μM primers (lane 3). (A, B) All reactions produced bands of the expected size (1614bp, arrowhead). (A) Additionally two smaller non-specific bands were obtained with approximate sizes of 250bp and 400bp in lanes 1 and 2, and thus, the product of lane 3 was digested with *Cla*I and *Xba*I and cloned into pCSHSP. (B) An additional non-specific by-product was observed in the reaction in lane 1 (cDNA prepared with random hexamers, white arrowhead). cDNA in lane 3 was also prepared with random hexamer primers. Reactions in lanes 2 and 4 (both cDNA prepared with oligo dT primer) were digested with *Cla*I and *Xba*I and cloned into pCSHSP. (C, D) PCR between primers 9A-1 and 9A-2, or 9B-1 and 9B-2 amplified the 1476bp *sox9a* or the 1390bp *sox9b* coding region (arrowhead), respectively, from plasmid minipreparation containing the entire *sox9a* or *sox9b* cDNA. Samples of reactions with 0.2μM primers (C, D lane 1), 0.5μM primers (C, D lane 2) and 1μM primers (C, D lane 3) are shown. (C) Reaction in lane 3 was digested with *Cla*I and *Xba*I and cloned into pCSHSP. (D) Products of lanes 1-3 were gel purified, digested with *Cla*I and *Xba*I and cloned into pCSHSP. In all panels, product sizes can be estimated using the quantitative 1kb DNA ladder (1kb M, NEB).


WT
1kb
M 1 2 3 M
1kb
2.0kb
1.5kb
1.0kb
0.5kb


B
1kb
m618
tw2
M 1 2 3 4
2.0kb
1.5kb
0.5kb
0.25kb


sox9a
1kb
M 1 2 3
3.0kb
2.0kb
1.5kb
1.0kb

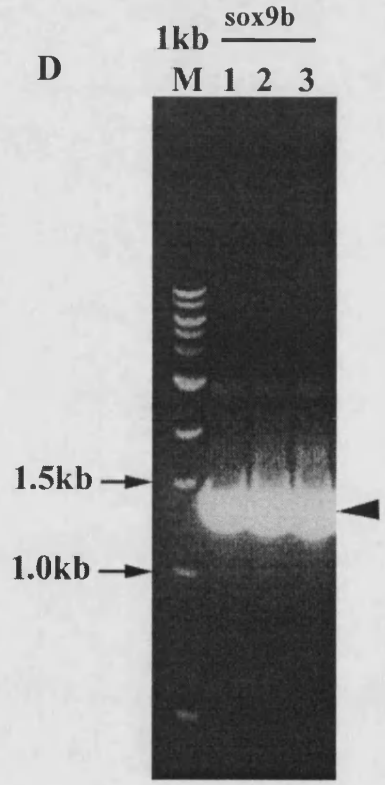
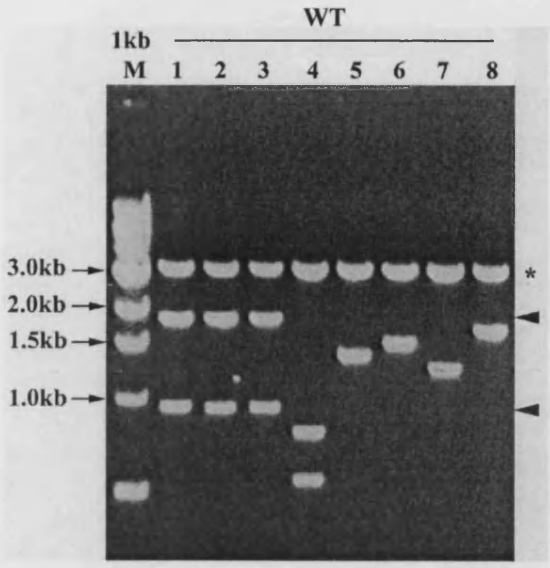
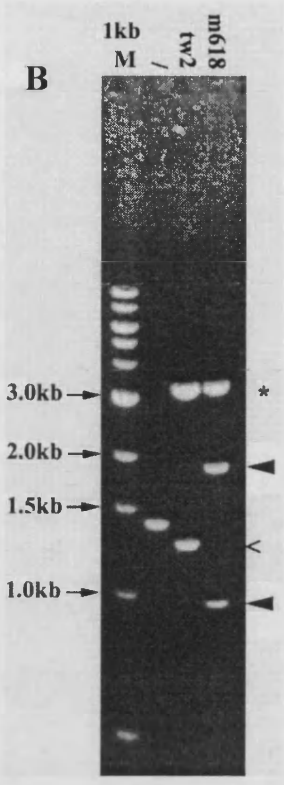
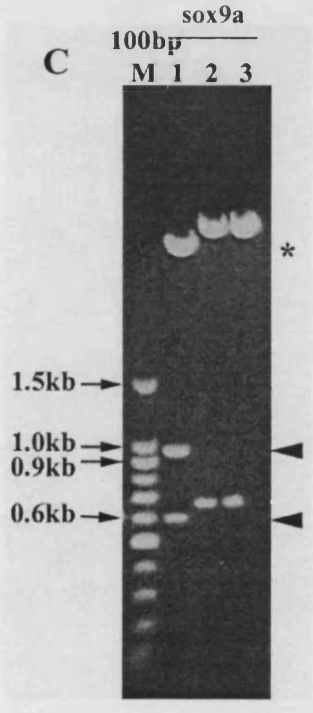
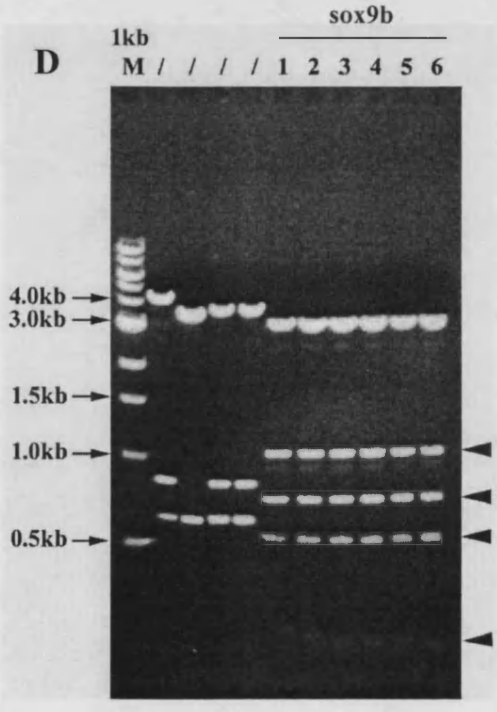

D
sox9b
1kb
M 1 2 3
1.5kb
1.0kb

**Figure 7.4:** Restriction digestion of putative heatshock constructs
(A, B) Samples of plasmid DNA minipreparations of putative hs>*cls*$^{WT}$ (A, lanes 1-8), hs>*cls*$^{m618}$ (B, lane m618) and hs>*cls*$^{tw2}$ (B, lane tw2) constructs were digested with *Pvu*II to test the size of the insert and are shown together with a quantitative 1kb DNA marker (1kb M, NEB). *sox10* coding regions contain a *Pvu* II restriction recognition site and thus, positive clones were expected to release a 967bp band and a 1871bp band (black arrowheads) in addition to a 3243bp band corresponding to the linearised vector (asterisk). Positive hs>*cls*$^{WT}$ clones (A, lanes 1-3) and one hs>*cls*$^{m618}$ clone (B, lane m618) could be identified, but the hs>*cls*$^{tw2}$ clone (lane tw2) was not correct, since the size of the insert (open arrowhead) was too small. The lane marked with a dash, /, is irrelevant to this experiment. (C) Samples of 3 putative hs>*sox9a* clones (lane 1-3) were digested with *EcoR*V and *Xba* I and positive clones were expected to show a 600bp and a 927bp band (arrowhead) compared to a 100bp marker (100bp M, Promega) together with a 3243bp band corresponding to the linearised vector (asterisk). (D) 6 samples of putative hs>*sox9b* (lane 1-6) were digested with *Pvu* II to test the size of the insert and shown here together with a quantitative 1kb marker (1kb M, NEB). Positive *sox9b* clones were expected to produce a restriction pattern consisting of a 1066bp, 809bp, 570bp and 213bp band (arrowheads) and a 3243bp band corresponding to the linearised vector (asterisk). (C, D) One positive clone each (C and D, lane 1) was confirmed by automated sequencing.


WT
1kb
M 1 2 3 4 5 6 7 8
3.0kb
2.0kb
1.5kb
1.0kb
*


B
1kb
M
tw2
m618
3.0kb
2.0kb
1.5kb
1.0kb
*


C
sox9a
100bp
M 1 2 3
1.5kb
1.0kb
0.9kb
0.6kb
*


D
sox9b
1kb
M / / / / 1 2 3 4 5 6
4.0kb
3.0kb
1.5kb
1.0kb
0.5kb

sequence when sequenced. None of 6 plasmid minipreparations derived from the *tw2* cloning released inserts of the correct size corresponding to the full length coding regions. However, sequencing a positive clone derived from the *m618* cloning revealed no aberrations to the wild-type protein sequence apart from the *m618* mutant lesion within the HMG domain. This construct will be referred to as hs>$cls^{m618}$ from this point onwards. Likewise, clones sox9A-1 and sox9B-1 proved to contain error free *sox9a* and *sox9b* coding regions respectively and will be referred to as hs>*sox9a* and hs>*sox9b* hereafter.

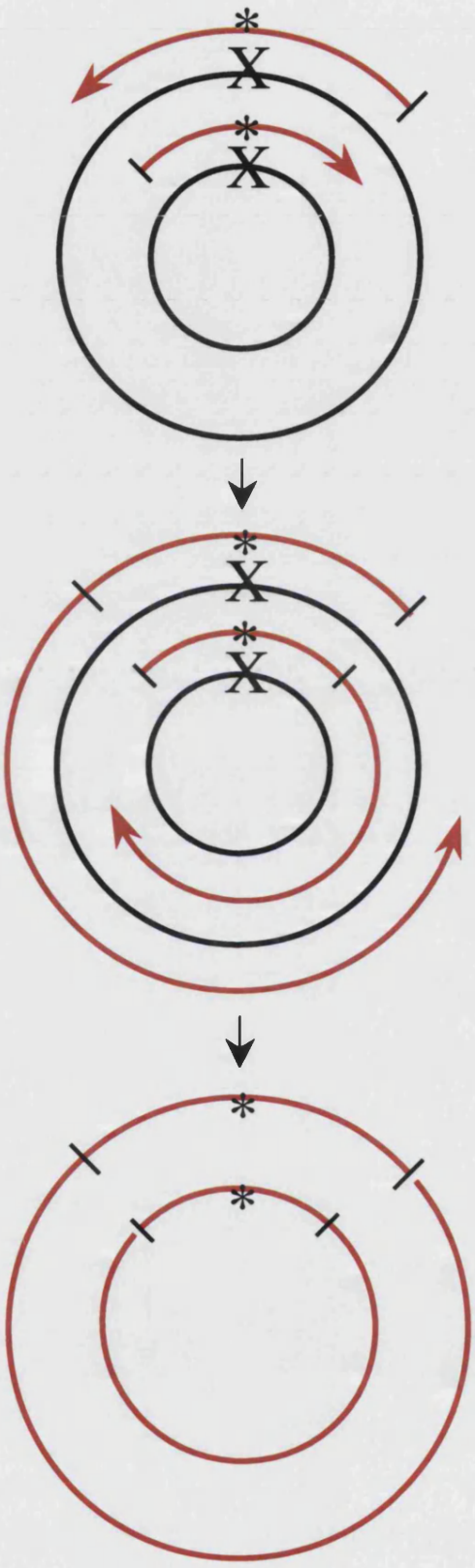
Since an error-free hs>$cls^{m618}$ construct was available and creating a wild-type (hs>$cls^{WT}$) and *tw2* (hs>$cls^{tw2}$) heatshock construct using RT-PCR had presented difficulties, a new strategy of site-directed mutagenesis of the hs>$cls^{m618}$ construct was chosen. We followed the guidelines from the QuikChange Site-Directed Mutagenesis Kit (Stratagene; Figure 7.5; see also Materials and Methods) to generate hs>$cls^{WT}$. Subsequently, the *tw2* mutant lesion was to be introduced into this wild-type construct by the same method to create the hs>$cls^{tw2}$ construct.

In the site-directed mutagenesis method, one or more nucleotide changes are introduced during *in vitro* replication of parental plasmid strands. The replication is primed from long reverse complementary mutagenesis primers, which contain the nucleotide mismatch to be introduced. To specifically remove the original parental strands after twelve rounds of strand replications, their bacterial origin and thus methylation status is utilized. In contrast to strands newly synthesised *in vitro*, parental strands derived from bacteria are methylated and are thus digested with the methylation sensitive restriction enzyme *Dpn* I. The remaining unmethylated, mutagenised linear DNA fragments align and form nicked open circular plasmids. After transformation into competent cells, these nicked plasmids are repaired and propagated normally.

Two pairs of mutagenesis primers were designed (Figure 7.6). The forward and reverse mutagenesis primers $cls^{m618}$→WT1 and $cls^{m618}$→WT2 encoded 35bp of wild-type sequence surrounding the *m618* nucleotide change. Thus, the nucleotide sequence of the resulting hs>$cls^{WT}$ would be identical to hs>$cls^{m618}$ apart from the T→A mutant lesion. Mutagenesis primers WT1→$cls^{tw2}$ and WT2→$cls^{tw2}$ were designed to introduce the *tw2* mutant lesion transforming hs>$cls^{WT}$ into hs>$cls^{tw2}$. The two constructs hs>$cls^{WT}$ and hs>$cls^{tw2}$ would be identical apart from the single change of nucleotide causing the K376X mutation.

**Figure 7.5:** Schematic of site-directed mutagenesis

The schematic was adapted and redrawn from the QuikChange Site-Directed Mutagenesis Kit Instruction Manual (Stratagene). The methylated parental strands (black) contain the site to be changed, indicated by an "X". Mutagenesis primers (red) contain the site to be introduced (asterisk) in the centre of the primer sequences. When parental strands are replicated, the mutation (asterisk) is incorporated. The parental strands (methylated, black) are selectively digested with a methylation sensitive restriction enzyme, whereas the new daughter strands (red) align and form open nicked circular plasmids, which are transformed into competent cells.



# Location of primers for site-directed mutagenesis

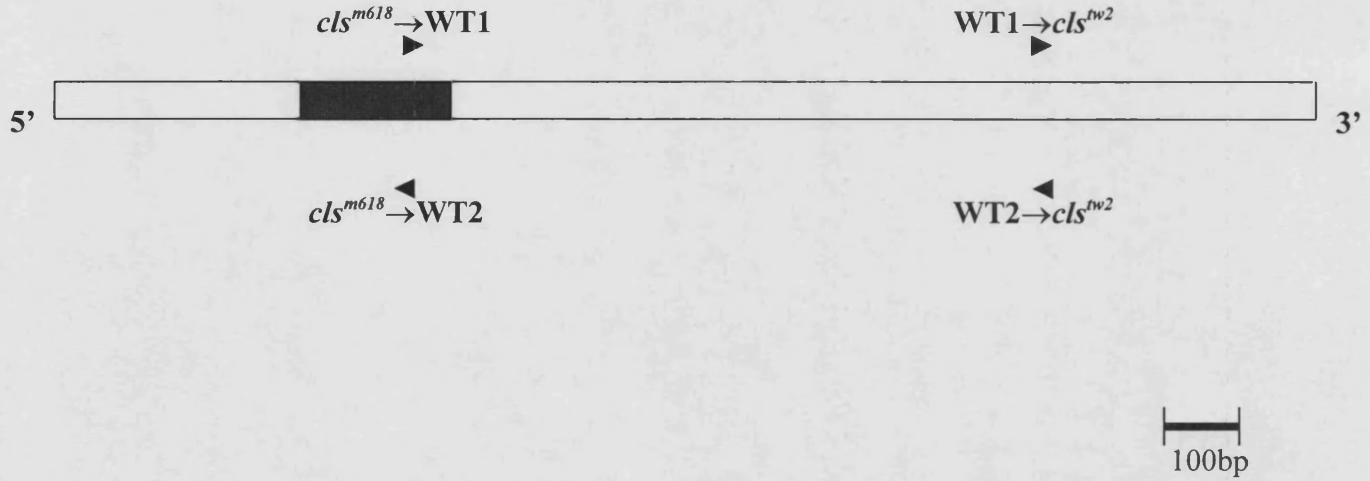
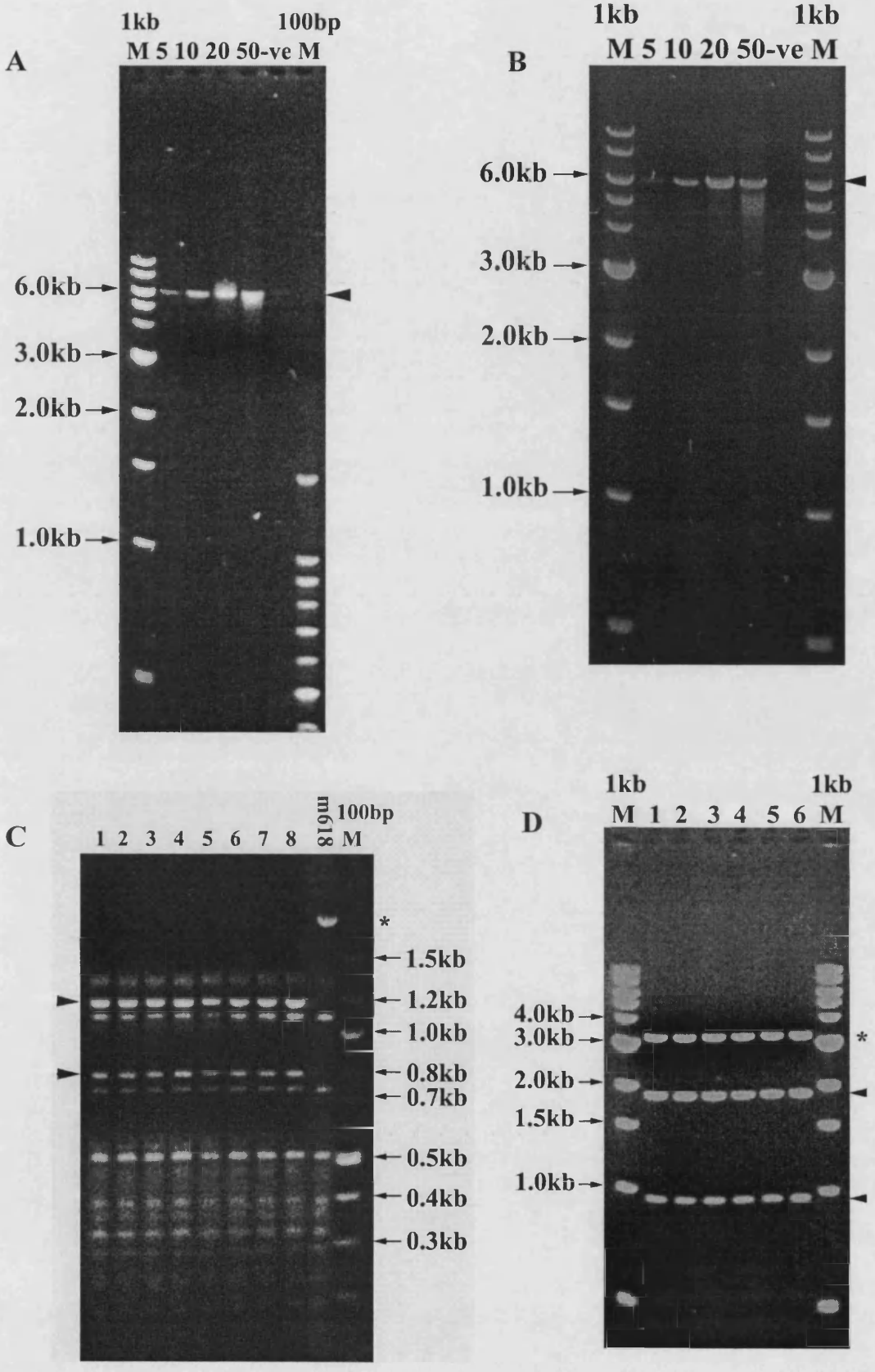



**Figure 7.6:** Mutagenesis primers to create the hs>$cls^{WT}$ and hs>$cls^{tw2}$ constructs

The *sox10* coding region is shown including the HMG domain (black box). The reverse complementary mutagenesis primers $cls^{m618}$→WT1 and $cls^{m618}$→WT2 were designed to transform $cls^{m618}$ into a wild-type coding region. Subsequently, the $cls^{tw2}$ mutation was introduced into the wild-type sequence using primers WT1→ $cls^{tw2}$ and WT2→ $cls^{tw2}$. The scale bar represents the length of a 100bp fragment.

Depending on which nucleotide change was to be introduced, reactions were set up with a range of amounts of either hs>$cls^{m618}$ or hs>$cls^{WT}$ template and a "no primer" control. After 12 cycles of replication, an aliquot of each reaction was analysed by agarose gel electrophoresis (Figure 7.7A, B). As expected, the amount of the 6103bp product generally increased with the amount of template. A very faint band was observed in the "no primer" control corresponding to the amount of template included in the reaction. However, a significant increase in product was observed in the equivalent reaction including the primers demonstrating successful replication. The reaction that produced the strongest product and the "no primer" control were separately digested with the methylation sensitive enzyme *Dpn* I to digest the parental methylated strands. A small aliquot was transformed into Gold Super-competent Epicurian Coli cells (Stratagene). As expected, no colonies were obtained from the control reaction due to complete digestion of parental methylated strands. In contrast, approximately 400 transformed colonies were counted from experimental reactions. 16 colonies from each of the two mutagenesis experiments were streaked out on fresh LB agar plates containing ampicillin to avoid genotypically mixed clones in a colony. A suitable number of these were picked and plasmid DNA isolated from them by plasmid minipreparations using the Wizard Plus SV Miniprep DNA Purification System (Promega). The potential hs>$cls^{WT}$ clones could be distinguished from parental hs>$cls^{m618}$ clones by a diagnostic restriction enzyme digest. In the wild-type sequence, an additional *Tsp*RI restriction site is created digesting an approximately 2000bp band, present in $cls^{m618}$, into a 1200bp and a 800bp band. Of 8 clones analysed all appeared to have lost the *m618* mutant lesion (Figure 7.7C). An equivalent diagnostic restriction digest was unavailable to test for successful synthesis of hs>$cls^{tw2}$ constructs. Thus, 6 clones were digested with *Pvu*II to test for inserts of the correct size. All 6 produced a pattern of 3 bands as expected for a *sox10* insert; a 3243bp vector band and 967bp and 1871bp bands, since the *sox10* coding region contains a *Pvu*II site (Figure 7.7D). One clone of each experiment, later renamed hs>$cls^{WT}$ and hs>$cls^{tw2}$ respectively, was sequenced to confirm successful mutagenesis.

**Figure 7.7:** Site-directed mutagenesis to create hs>*cls*$^{WT}$ and hs>*cls*$^{tw2}$
(A, B) Samples of the site-directed mutagenesis reaction converting hs>*cls*$^{m618}$ into hs>*cls*$^{WT}$ (A) and hs>*cls*$^{WT}$ into hs>*cls*$^{tw2}$ (B) are shown together with quantitative 1kb markers (1kb M, NEB). Experimental reactions were set up with 5ng (lane 5), 10ng (lane 10), 20ng (lane 20), and 50ng (lane 50) template. A negative "no primer" control (-ve) was included with 20ng template. As expected, the amount of the 6103bp product (arrowhead) generally increased with the amount of template. A very faint band was observed in the "no primer" control corresponding to the amount of template included in the reaction. (C, D) An aliquot of the reaction from lane 50 (A) and from lane 20 (B) was transformed into Gold Super-competent Epicurian Coli cells (Stratagene). (C) 8 clones (lanes 1-8) were restriction digested with *Tsp*RI to distinguish between parental hs>*cls*$^{m618}$ (lane m618) and successfully created hs>*cls*$^{WT}$ clones. In the latter wild-type clones, a *Tsp*RI site was expected to be gained resulting in an approximately 1.2kb and a 800bp band (arrowheads) instead of a 2.0kb band seen in m618 (asterisk). All 8 clones tested had successfully lost the m618 mutant lesion. A quantitative 100bp marker (100bp M, NEB) was loaded for size comparison. (D) 6 clones (lanes 1-6, D) were restriction digested with *Pvu*II to test for inserts of the correct size corresponding to wild-type or mutant *sox10* coding regions, since no diagnostic restriction digestion for successfully introduced *tw2* mutant lesions was available. All 6 clones produced the expected bands, 1871bp and 967bp (arrowheads) and a 3243bp band (asterisk) corresponding to the linearised vector. A quantitative 1kb marker (1kb M, NEB) was loaded for size comparison. (C, D) One positive clone each (C, lane 1 and D, lane 4) was confirmed by sequencing and renamed hs>clsWT and hs>*cls*$^{tw2}$, respectively.


A
1kb
100bp
M 5 10 20 50-ve M
6.0kb
3.0kb
2.0kb
1.0kb
B
1kb
1kb
M 5 10 20 50-ve M
6.0kb
3.0kb
2.0kb
1.0kb
C
1 2 3 4 5 6 7 8
mG18
100bp
M
*
1.5kb
1.2kb
1.0kb
0.8kb
0.7kb
0.5kb
0.4kb
0.3kb
D
1kb
1kb
M 1 2 3 4 5 6 M
4.0kb
3.0kb
2.0kb
1.5kb
1.0kb
*

## Injection of the hs>*cls*$^{WT}$ construct shows partial rescue of the *cls* phenotype

The following experiments were performed to test whether ectopic *sox10* driven by a heatshock promoter could rescue the *cls* mutant phenotype.

Firstly, we injected a suitable concentration of the hs>*cls*$^{WT}$ construct into the yolk of early cleavage stage embryos from a *cls*$^{m618}$ heterozygous cross (Materials and Methods). At this early stage, the construct is distributed to all blastomeres by yolk streaming, since there are no physical boundaries between blastomeres and yolk up to the 8 cell stage (Kimmel et al., 1995). The embryos were then raised at 28.5°C. In initial experiments, based on descriptions of rescue with a *nacre* heatshock construct (Lister et al., 1999), two one hour heatshocks were carried out at 37°C, the first between 10-13hpf and the second heatshock between 22-24hpf. The rationale for choosing these time points was that *sox10* expression is first observed at the 2 somite stage corresponding to 10 ½ hpf. Secondly, the first phenotype in *cls* embryos is apparent at approximately 24hpf, when pigment cell precursors on the lateral pathway fail to migrate (Kelsh et al., 1996). Thus, Sox10 function commences before or at approximately this time. Embryos from a *cls*$^{m618}$ cross were chosen, because preliminary phenotypic studies (previous chapter) indicated that it might be one of the strongest alleles and for reasons of consistency since all *in situ* hybridisations had been carried out on *cls*$^{m618}$ embryos.

Of a total of 92 injected *cls* mutant embryos that survived, 44 (48%) were rescued. On average, they developed 15 melanophores per embryo with stellate wild-type morphology and dark pigmentation (Figure 7.8, Appendix 7.1A).

### *Investigation of deformities*

Although hs>*cls*$^{WT}$ clearly rescued the melanophore phenotype in *cls* mutants, 63.8% of injected embryos also showed various degrees of malformations. Defects ranged from slightly smaller eyes and mildly deformed axes to grossly amorphic tissue with pigment cells surrounding the yolk sac.

We were concerned about malformations since they might have an influence on the rescue experiments. For example, to identify a rescued embryo, its eyes have to be present and fully pigmented in order to distinguish it from a malformed wild-type sibling. A developmentally delayed wild-type sibling might easily be mistaken for a

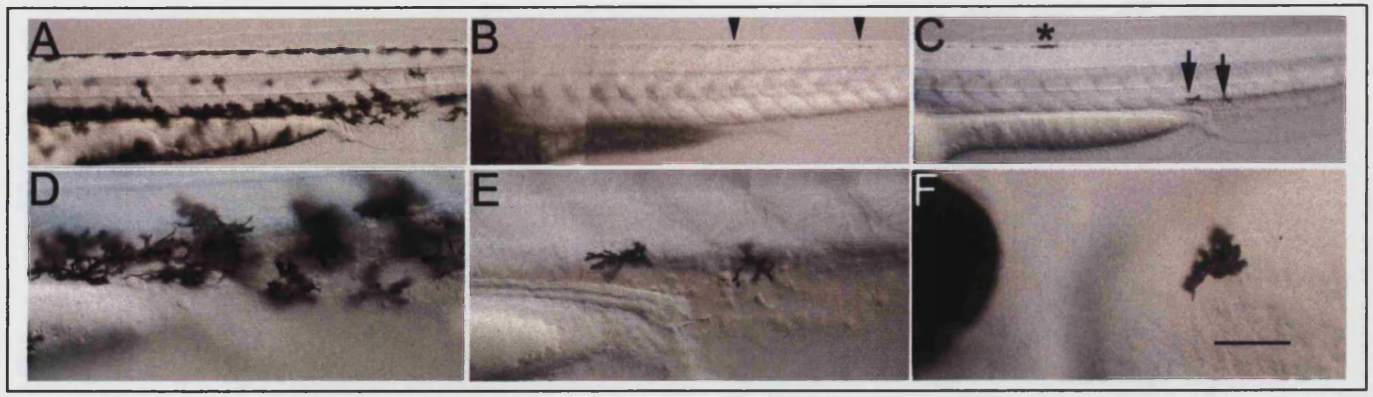



**Figure 7.8:** Rescue of the pigment phenotype in homozygous *cls* mutants by ectopic expression of hs>*cls*$^{WT}$

(A-C) In 48 hpf wild-type embryos, numerous neural crest precursors have differentiated into stellate, pigmented melanophores (A). Such cells are never observed in uninjected homozygous *cls*$^{m618}$ mutants (data not shown) or when injected with hs>*cls*$^{m618}$ (B). Occasionally, faintly pigmented, rounded cells are seen in the premigratory neural crest area (B, arrowheads). After ectopic expression of hs>*cls*$^{WT}$, a variable number of melanophores are rescued to a wild-type morphology and are able to migrate to ventral locations (C). In this embryo, one melanophore was noted in the premigratory neural crest area (C, asterisk), one on the yolk sac (F) and two in the anal region (C, arrows). Rescued melanophores adopt a morphology characteristic for their final location as can be seen by comparing the melanophores in the anal region (arrows in C, enlarged in E) with the equivalent region in wild-type siblings (D). Scale bar: 125μm (A-C), 70μm (D-F).

rescued *cls* embryos, but this can be overcome by comparing the extent of pigmentation in its eye to other wild-type siblings. An undoubtedly rescued *cls* embryo would show the same extent of pigmentation in the eye as its wild-type siblings, but less along its body.

Furthermore, heavily malformed embryos might not show any rescue or rescue to a lesser degree regardless of Sox10's functional abilities.

A low level of these deformities were likely to be caused by a proportion of bad eggs and/or by injection trauma. To control for bad batches of eggs, we also raised 10-30 uninjected siblings from each batch and scored the proportion of surviving embryos (Appendix 7.1) and the percentage of malformed within the surviving embryos (Appendix 7.3). The difference between the proportion of deformed embryos in uninjected and injected siblings was a measure for malformation due to the injection process, overexpression of Sox10 protein or toxicity of the heatshock plasmid itself. Trauma caused by the injection process likely varied slightly on a daily basis due to personal skilfulness and differences in injection needles. However, by comparing the mean results of several batches, we hoped to control for this variability.

To investigate the origin of these deformities and optimise the injection protocol accordingly, we carried out a series of experiments testing the effect of parameters likely to influence the occurrence of malformation such as the amount of construct injected and the time of heatshock.

First, we carried out a $\chi^2$-test (using the Yates' correction formula) on *cls*/WT pairs of individual datasets to test whether these malformations were linked to a certain genotype. It was asked whether the ratio of malformed:total wild-type (WT) embryos was different to the equivalent ratio in *cls* embryos. Appendix 7.3 lists all malformation data collected. Under all conditions, no significant difference was found in malformation frequency between *cls* mutant and wild-type embryos (Appendix 7.4A). We then employed a single factor ANOVAR analysis with 2 levels to investigate the effect of various parameters, such as heatshock (hs) treatments, on the frequency of malformed embryos.

First, we asked, whether the heatshock treatment itself increased the amount of malformations. To do so, we compared the number of malformed embryos depending on early (15-16hpf) or late hs (18-19.5hpf) versus no hs at both 50pg (high) and 25pg (low) DNA injected per embryo. The "no hs" control was included to assess the leakiness of the heatshock promoter. All experiments were scored for percentage of

malformed embryos (Appendix 7.3). At high concentration, a very significant difference at the 0.1% level was found regardless of whether it was an early or late heatshock, whereas at low concentration, no significant difference was observed even at the 5% level (Appendix 7.4B). Thus, an early or late heatshock does increase malformations, but only when high concentrations of DNA are injected.

Next, we asked whether the toxic effect observed was due to the *sox10* cDNA or the plasmid vector itself. We injected embryos with 50pg (high) and 25pg (low) pCS plasmid only (without the *sox10* cDNA) and heatshocked early and late. We found no significant difference at the 0.1%, 1% or 5% level between pCS or hs>$cls^{WT}$ injections when constructs were injected at the low concentration, irrespective of the time of heatshock and at the high concentration when heatshocked late. When high concentrations were injected and embryos heatshocked early, there seemed to be a significant difference between pCS and hs>$cls^{WT}$ at the 5% level. This could point to a slight toxicity of *sox10* when overexpressed during early development. For all the other conditions, the malformations appeared to be attributable to toxic effects of the heatshock plasmid alone (Appendix 7.4C).

When we compared the effect of high and low amounts of hs>$cls^{WT}$ construct on malformation, no significant difference was found when heatshocked early, but a very significant difference (at the 0.1% level) when heatshocked late. The equivalent comparisons with pCS injections resulted in a significant difference at the 1% level between high and low concentration regardless of the time of heatshock (Appendix 7.4D).

Next, we investigated whether the time of heatshock had an effect on the percentage of deformed embryos. Consistent with the heatshock versus no heatshock results, a significant difference between early and late hs was observed when high concentrations of either hs>$cls^{WT}$ or pCS was injected, but no significant difference at low concentration. Furthermore, the effect was more pronounced in the presence of *sox10* cDNA (Appendix 7.4E).

Finally, we asked whether malformations had any effect on the rescue experiment. For example, are rescued embryos more or less likely to be malformed than others with normal morphology. We tested this with a $\chi^2$-test and investigated whether the ratio normal:malformed within the rescued *cls* embryo population was significantly different from 1:1. Individual tests were carried out within the different types of treatments and also for all of the 22 datasets together (Appendix 7.4F). In any case,

within the population of rescued embryos, the number of malformed embryos was not significantly higher than rescued embryos with normal morphology. Thus, the presence of malformation did not appear to affect the proportion of *cls* embryos rescued.

The results of these previous experiments generally indicated that malformations tended to be lower if less DNA was injected and the hs was carried out earlier, at 15-16hpf. Thus, to optimise the assay further, we tested an even lower (12pg) and higher (120pg) concentration of hs>$cls^{WT}$ in combination with an early hs. For each of the 4 construct concentrations, embryos were scored for the percentage of rescued *cls* mutants, the degree of rescue (mean number of melanophores per rescued embryo) and the percentage of malformation. For comparison, all results were combined and plotted in a graph (Figure 7.9, Appendix 7.1A). As expected, the percentage of rescued *cls* embryos and malformations generally increased with the amount of DNA injected. However, the number of rescues seemed to plateau or slightly decrease between 50pg and 120pg of hs>$cls^{WT}$, whereas the deformities steadily increased. The mean number of melanophores remained fairly constant throughout, apart from a significant peak of 3 times increase at 50pg.

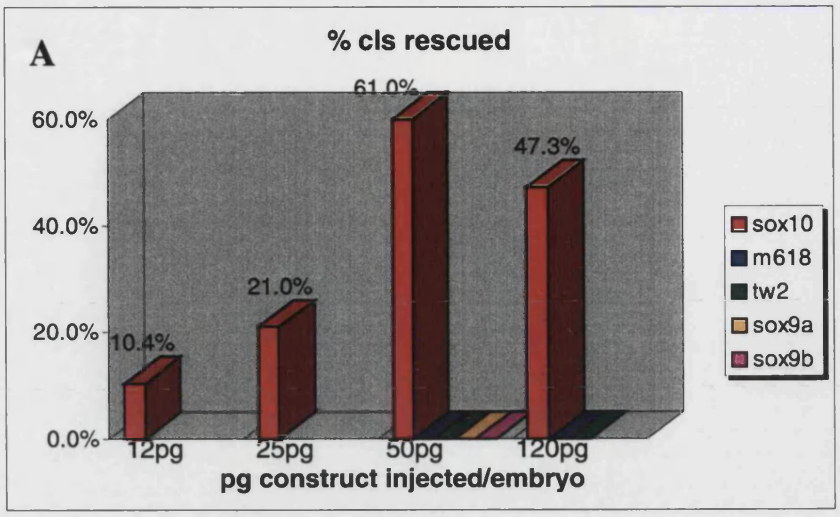
Taken together, in our hands 50pg construct in combination with an early hs at 15-16hpf produced the highest degree of rescue (on average 18.7 melanophores) in the highest proportion (61%) of *cls* mutant embryos with relatively low levels of malformation (35.7%).

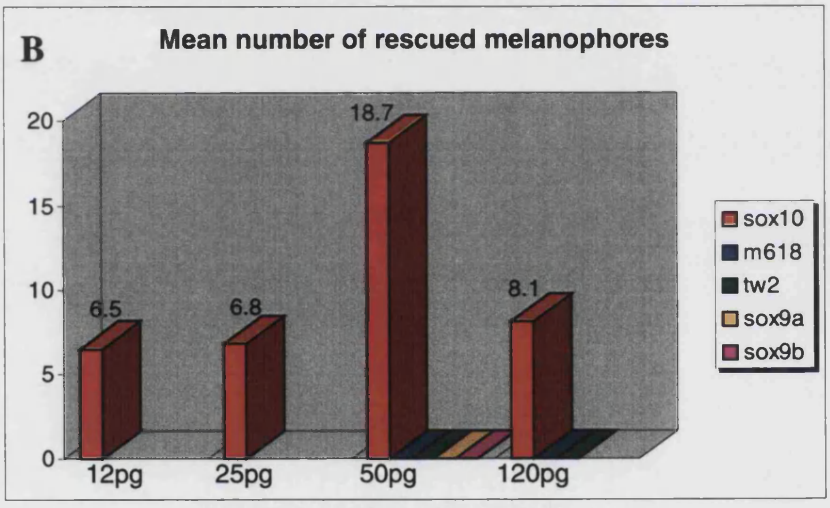
## Can mutant *sox10* alleles rescue the melanophore aspect of the *cls* phenotype?

In the previous section we could show that hs>$cls^{WT}$ was able to partially rescue the melanophore phenotype in *cls* mutant embryos. We decided to use the same assay to characterise some of the *cls* mutant alleles in an *in vivo* system. We wanted to ask whether the mutant proteins had retained any residual function by testing the ability of them to rescue pigment cells. Furthermore, we were interested to see whether alleles exhibited any dominant negative effects as had been hypothesised for some of the human *Sox10* alleles (1400del12, Inoue et al., 1999; E189X and 1076delGA, Pingault et al., 1998a) and so far only been shown conclusively *in vitro* for the mouse allele C190X (*Sox10del*, Potterf et al., 2000). In accordance with the mammalian *Sox10* pigmentation phenotype, we predicted that such a dominant

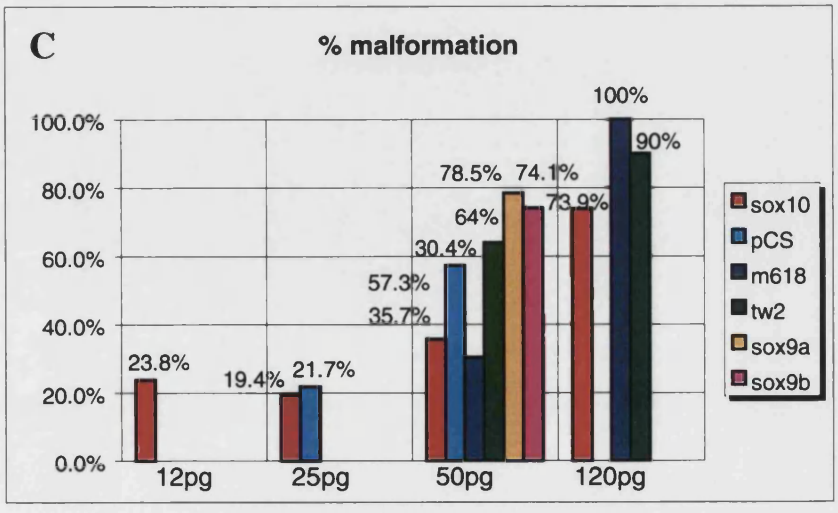
**Figure 7.9:** Summary of rescue and malformation data

(A, B) The graphs show the percentage of homozygous *cls* mutant embryos rescued after injection of various amounts of either hs>*cls*$^{WT}$ (sox10), hs>*cls*$^{m618}$ (m618), hs>*cls*$^{tw2}$ (tw2), hs>*sox9a* (sox9a) or hs>*sox9b* (sox9b) (A) and the mean number of melanophores per rescued *cls* embryo (B). hs>*cls*$^{m618}$ and hs>*cls*$^{tw2}$ were only assessed at 50pg and 120pg, whereas hs>*sox9a* and hs>*sox9b* only at 50pg. Only hs>*cls*$^{WT}$ was able to rescue the pigment phenotype in *cls* embryos. The highest percentage rescue and rescued melanophores per *cls* embryo were observed with 50pg. (C) The graph depicts the percentage of all malformed embryos (wild-type and *cls* mutants) observed for each amount of construct injected. The same abbreviations were used as described in (A) with the addition of the heatshock vector without the *sox10* coding region (pCS). Regardless of the construct, malformations increased with the amount of construct injected. hs>*sox9a* and hs>*sox9b* were only assessed at 50pg, pCS only at 25pg and 50pg and hs>*cls*$^{m618}$ and hs>*cls*$^{tw2}$ were tested at 50pg and 120pg.




pg construct injected/embryo




pg construct injected/embryo

negative effect in zebrafish would be reflected in a reduction in number, aberrant distribution or perhaps changed morphology of melanophores due to cell death. No phenotype was noticed in uninjected heterozygous siblings of a heterozygous *cls*$^{m618}$ or *cls*$^{tw2}$ cross. However, the dominant effect might only become apparent when mutant *sox10* is overexpressed in wild-types or in heterozygous mutants, in which the amount of Sox10 is halved already.

Consistent with the injection protocol for hs>*cls*$^{WT}$, in most cases an early heatshock at 15-16hpf was performed. As a control for the initial experiment with hs>*cls*$^{WT}$, most hs>*cls*$^{m618}$ injected embryos were also subjected to two heatshocks at the time points used before. There was the possibility that a mutant protein with only little residual function would be required at higher concentrations for rescue than the wild-type protein. Thus, 50pg and 120pg of hs>*cls*$^{m618}$ and hs>*cls*$^{tw2}$ was injected as described before. Even though we expected to see a high percentage of malformations as a result, the chances of rescue might be increased as well.

Of a total of 125 *cls*$^{m618}$ mutant embryos injected with either 50pg or 120pg hs>*cls*$^{m618}$ that survived to be scored for rescue, none showed any melanophores with wild-type morphology (Figure 7.9, Appendix 7.1B). Uninjected siblings usually showed only very low levels of deformities. Although the proportion of malformations in the uninjected siblings were comparable to the equivalent 50pg hs>*cls*$^{WT}$ injections, indicating a batch of eggs with similar quality, the percentage of malformed embryos injected with hs>*cls*$^{m618}$ was approximately half of that of the hs>*cls*$^{WT}$ injections (Figure 7.9, Appendix 7.3). When 120pg hs>*cls*$^{WT}$ or hs>*cls*$^{m618}$ was injected, the proportion of malformed embryos was comparable.

Similarly, none of 249 *cls*$^{tw2}$ embryos could be scored as rescued (Figure 7.9, Appendix 7.1C). In these injections, 3 embryos injected with 120pg hs>*cls*$^{tw2}$ and 3 injected with 50pg exhibited 1-12 melanophores in the dorsal and ventral stripe of trunk and tail. Unfortunately, all 6 were very severely malformed and lacked eyes or even head structures, which made it impossible to identify them as *cls* mutant embryos. We could not exclude the possibility of them being severely delayed wild-type siblings and thus, they could not be counted as rescues. Although comparable proportions of deformed embryos were recorded in uninjected siblings, twice as many embryos injected with 50pg of hs>*cls*$^{tw2}$ were malformed than with hs>*cls*$^{WT}$, whereas with 120pg the percentages were not significantly different (Figure 7.9, Appendix 7.3).

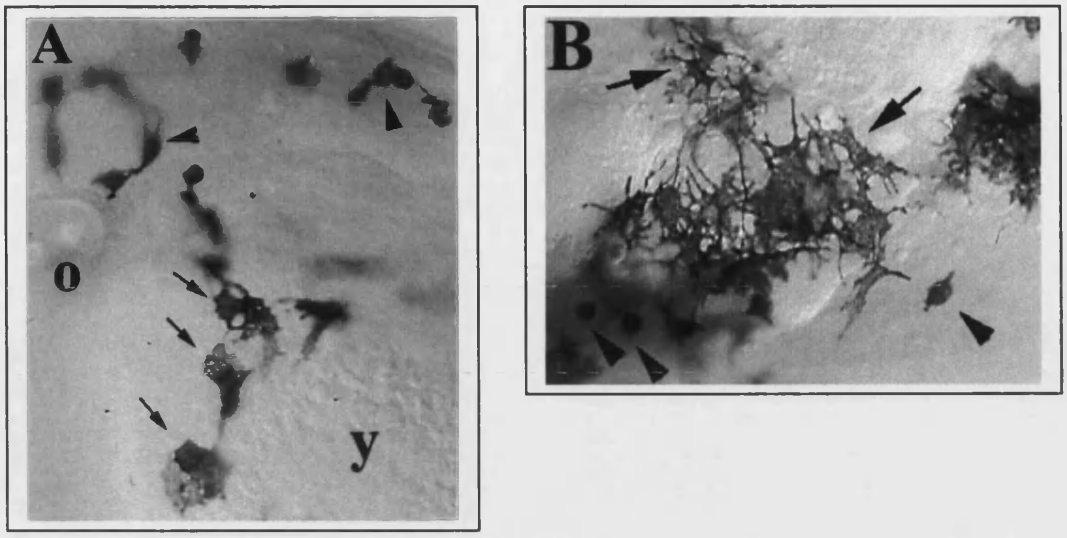
We noticed melanophores with abnormal morphology in 20-90% of phenotypically wild-type embryos obtained from a heterozygous cross that had been injected with hs>*cls*$^{tw2}$ (Figure 7.10A, B). In the majority of cases, the percentage was between 55-85% (Figure 7.10C). As described in the beginning of this section an abnormal melanophore morphology might be expected for a dominant negatively acting Sox10 protein. Furthermore, the proportion of embryos exhibiting this phenotype was also broadly consistent with wild-type and heterozygous mutant siblings, in which a possible dominant negative mutant protein had been overexpressed. Thus, a possible dominant negative function of *cls*$^{tw2}$ was further investigated. A simple genetic test eliminated this hypothesis very quickly. If the *cls*$^{tw2}$ allele encoded a dominant negative protein, a similar pigmentation phenotype would be expected to be seen in all heterozygotes of an uninjected *cls*$^{tw2}$ cross. No abnormal melanophores were observed in heterozygous *cls* mutants, thus ruling out any dominant negative effect of this allele at least on the pigment cell lineage.

## Can *sox9a* or *sox9b* rescue the melanophore aspect of the *cls* phenotype?

For reasons explained in the introduction to this chapter, we were interested to investigate a possible functional redundancy between *sox10* and its closest relative, *sox9*. The same rescue protocol and scoring criteria were applied as had been established for the rescue experiments with hs>*cls*$^{WT}$.

Of 325 surviving embryos after injection with 50pg hs>*sox9a* and heatshocked at 15-16hpf (early), 76 *cls* embryos were identified. None exhibited any rescued melanophores (Figure 7.9, Appendix 7.2A). Although the level of malformation in uninjected siblings was comparable, injection with 50pg of hs>*sox9a* instead of hs>*cls*$^{WT}$ appeared to result in an increased proportion of deformed embryos (Figure 7.9, Appendix 7.3).

A similar result was obtained with hs>*sox9b*. Of 263 surviving embryos injected with 50pg hs>*sox9b* and heatshocked early, none of the 58 *cls* embryos were rescued (Figure 7.9, Appendix 7.2B). Again, embryos injected with hs>*sox9b* showed a higher proportion of malformed embryos than a comparable batch injected with hs>*cls*$^{WT}$ (Figure 7.9, Appendix 7.3).




C

| Injection batch | Abn. M/total embryos | % Abn. M |
|---|---|---|
| T-4 | 4/16 | 25% |
| T-5 | 16/82 | 19.5% |
| T-6 | 74/130 | 56.9% |
| T-7 | 37/67 | 55.2% |
| T-8 | 16/39 | 41% |
| T-9 | 19/28 | 67.9% |
| T-10 | 16/19 | 84.2% |

**Figure 7.10:** Investigation of a possible dominant negative effect in *cls*$^{tw2}$

(A, B) The two panels show examples of the aberrant melanophore phenotype (black arrowheads) in 48hpf embryos of a heterozygous *cls*$^{m618}$ cross injected with the hs>*cls*$^{tw2}$ construct, which could indicate a possible dominant negative (DN) function of the *cls*$^{tw2}$ allele. These melanophores appeared rounded, often bipolar or elongated and were dispersed between melanophores with wild-type morphology (black arrows). o, otic vesicle; y, yolk sac. (C) The table summarises counts of embryos in which such a phenotype was observed. The first column indicates the injected batch analysed, column 2 contains the number of embryos exhibiting the abnormal melanophores (Abn. M) in relation to the total number of embryos injected in this batch (total embryos). In column 3, the percentage of embryos that show these abnormal melanophores is given (% Abn. M). Genetically, ¾ (75%) of embryos corresponding to wild-type and heterozygous *cls*$^{m618}$ siblings, might exhibit the aberrant phenotype, but at least all heterozygotes (50%), in which the amount of Sox10 is halved already.

## Discussion

We have described the successful rescue of the pigment phenotype in *cls*$^{m618}$ mutants by ectopic expression of *sox10*. This was achieved by driving the expression of *sox10* with a heatshock inducible promoter. In addition, we demonstrated that neither Sox9a nor Sox9b can replace Sox10 function in this assay.

Generation of the hs>*cls*$^{WT}$ construct by RT-PCR and cloning into the heatshock vector was unsuccessful. Even though a high-fidelity Taq polymerase was used in this PCR reaction, it seemed difficult to identify such a hs>*cls*$^{WT}$ construct without any non-conservative nucleotide changes, which altered the amino acid sequence. The reason for this increased error rate might have been the high number of PCR cycles required to amplify the coding region from cDNA. However, an error-free hs>*cls*$^{m618}$ was identified and thus, using site-directed mutagenesis to create the missing heatshock constructs had several advantages. In this method, all mutagenised strands are created by a single replication of only parental strands (Figure 7.5) and thus, any mistake introduced is not amplified in subsequent rounds of replication. Hence, this approach proved to be an easy and very reliable method, which was employed repeatedly to create 2 heatshock constructs. The sample of constructs analysed all proved to have incorporated the desired nucleotide change and were tested in *in vivo* rescue assays.

Injection of 50pg hs>*cls*$^{WT}$ construct followed by a heatshock treatment resulted in rescue of a mean number of 15 melanophores in 48% of surviving *cls* mutant embryos. In the control experiment, in which embryos were also injected with 50pg of hs>*cls*$^{WT}$ but not heatshocked, resulted in only a slightly lower percentage of rescued embryos, 35% instead of 48%. This indicated a fairly high degree of leakiness of the heatshock promoter. *sox10* transcripts were thus produced even without the requirement of a heatshock. This finding did not invalidate the rescue results. Firstly, although the number of rescued embryos was similar with or without heatshock, the proportion of rescue was significantly higher in heatshocked embryos (mean number of rescued melanophores per embryo 15.0 instead of 4.3). Secondly, embryos injected with the control construct hs>pCS, failed to rescue melanophores, whether or not heatshocked. Thirdly, a similarly high degree of leakiness for the same heatshock plasmid was reported for *nacre* rescue experiments (Lister et al., 1999).

The leakiness of the heatshock promoter might also be partially responsible for the high percentage of malformation observed in injected embryos. However, our control

injections revealed that a large proportion of the deformations observed were not likely to be caused by Sox10, but by either the plasmid itself or the heatshock promoter. The proportion of malformed embryos obtained with hs>$cls^{WT}$ was usually not significantly different to those obtained with an equivalent amount of control plasmid hs>pCS, which only lacked the *sox10* sequence. Additionally, the optimisation procedure revealed that not only the proportion of embryos with malformations, but also the percentage and degree of melanophore rescue, generally increased with increasing amounts of *sox10* DNA injected. The former result was consistent with the construct showing slightly toxic effects. Furthermore, one instead of two heatshock treatments lowered the proportion of deformed embryos, although a heatshock treatment in general did not increase malformations as long as low amounts of construct were injected. This could be interpreted as a slight toxicity of the Sox10 protein itself. Lastly, if the heatshock was carried out early (15-16hpf) instead of late (18-19.5hpf), slightly less malformed embryos were observed. The reason for this might be greater sensitivity to high levels of Sox10 at earlier stages during development. Misregulation of gene transcription at this earlier stage is likely to result in more deleterious defects.

In summary, the best results overall in our assay were achieved with 50pg hs>$cls^{WT}$ and a heatshock at 15-16hpf. The amount of construct injected per embryo is similar to that reported for rescue of homozygous *nac* embryos with hs>*nac* (Lister et al., 1999). However, more *nacre* embryos were rescued if the heatshock treatment was carried out later, between 18-20hpf.

For future experiments, it might be worthwhile to assess whether the toxic effects observed above resided in the plasmid or the heatshock promoter itself. This could be tested by comparing deformities after injection of the expression plasmid with and without the heatshock promoter. In the case of promoter toxicity, rescue experiments might be much improved if Sox10 expression was regulated by another promoter, e.g. its own promoter, especially if experiments require injection of large amounts of construct as in the case for testing mutant alleles. The *sox10* promoter would also direct Sox10 expression to more specific locations and time.

Rescued melanophores were observed mosaicly as was expected for DNA injections into early embryos (Oliver et al., 1996; Hammerschmidt et al., 1999). During development, hs>$cls^{WT}$ is distributed unevenly between dividing daughter cells and thus, not every cell will have received the construct.

Sox10 not only enabled pigment precursors to differentiate and adopt the typical morphology for wild-type melanophores, but also rescued their migration to ventral positions. Thus, rescued melanophores were often observed in the ventral stripe or on the yolk sac. The role of Sox10 could therefore be primarily in enabling the migration of crest cells, which then encounter the appropriate signals for differentiation. Judging by the *sox10* expression pattern in *cls* embryos at approximately 24-30hpf, a greater number of *sox10*-positive neural crest cells failed to migrate and appeared to remain in the crest staging area compared to wild-type. They still did not differentiate in that location as would be expected from a primarily migrational defect. Furthermore, the premigratory area is also the final location for some melanophores residing in the dorsal stripe and hence all necessary factors for differentiation have to be present. However, melanophores still did not differentiate properly in *cls* mutants in this area. In summary, the most likely explanation for the observations above is that neural crest cells that receive sox10 are able to become specified as pigment cell precursors and so are enabled to migrate to their final locations.

The rescue of *cls* mutants by heatshock after injection of a hs>$cls^{WT}$ construct raised the question of whether mRNA injection might also be an appropriate strategy for achieving rescue. However, melanophores failed to be rescued by *sox10* mRNA injection (S. Elworthy, unpubl. data). This result was not surprising since Sox10 may not be required until approximately 24hpf of development, whilst it is unlikely that injected RNA is still present in sufficient amounts this late (Hammerschmidt et al., 1999). However, the injected *sox10* mRNA was functional, since transcripts of a downstream target, *nacre*, were detectable by *in situ* hybridisation in 6hpf stage embryos, some 12 hours prior to endogenous gene expression (Lister et al., 1999).

It will be very interesting to test for other rescued neural crest derivatives in future experiments. This has been attempted for the enteric lineage (S. Elworthy and R. Kelsh, unpubl. data), but to date has been unsuccessful, perhaps because the number of enteric precursors maybe very small compared to pigment precursors. Due to the mosaic characteristics of the injection technique, this small group might be very difficult to target and thus, rescue would be much harder to observe.

In $cls^{m618}$, the L142Q lesion is expected to significantly decrease the binding ability of Sox10 to its recognition sequences in the promoter of downstream targets. Reduced DNA binding ability might also prevent $Sox10^{m618}$ from interacting with its partner transcription factors, since both the HMG and the synergy domain are required

for synergistic activity (Kuhlbrodt et al., 1998a). Furthermore, failure to interact synergistically with other transcription factors was observed for similar human mutations containing disrupted HMG domains (Kuhlbrodt et al., 1998b). Diminished DNA binding might also affect the efficiency of the transactivation domain as was demonstrated for the human 482ins6 mutant (Bondurand et al., 2000). Although the $cls^{tw2}$ allele has likely retained its binding ability, it has lost the transactivation domain due to a premature truncation of the protein caused by the K376X lesion. Consistent with this, these two *cls* alleles failed to show any significant transcriptional activation abilities in the *in vitro* study described in the previous chapter, indicative of them being null alleles. However, there is a possibility that they might show some residual function through their putative synergy domains. This is less likely to be true for $cls^{m618}$, since a functional DNA binding domain is required for synergistic activity, at least for rat Sox10 (Kuhlbrodt et al., 1998a). Thus, taken together, it seemed unlikely that either of the two alleles, but in particular $cls^{m618}$, would rescue any aspect of the *cls* phenotype. As predicted, both mutant Sox10 proteins encoded by $cls^{m618}$ and $cls^{tw2}$ failed to rescue the pigment aspect of the *cls* phenotype. It would be interesting to investigate whether other neural crest derivatives affected in *cls* embryos, for example the peripheral nervous system (PNS), could be rescued.

All human and mouse alleles exhibit a dominant phenotype. Some of those dominant effects are thought to be due to dominant negativity and manifest themselves in patients with white forelocks and occasionally depigmented skin patches, blue eyes or heterochromic irides (Pingault et al., 1998a; Bondurand et al., 1999; Southard-Smith et al., 1999) and white belly spots and feet in mice (Lane and Liu, 1984; Herbarth et al., 1998; Southard-Smith et al., 1998). In batches of embryos injected with $cls^{tw2}$, melanophores with abnormal morphology were observed, which in principle could have been due to a dominant negative effect. However, these abnormal melanophores were not seen in heterozygous offspring of an uninjected $cls^{tw2}$ cross. Furthermore, petri-dishes and embryo medium had irrationally been re-used once for these batches of injected embryos. Thus, possible changes of the medium concentration due to evaporation, perhaps exacerbated by injection trauma, are more likely to have caused this aberrant pigment morphology.

We have created two further heatshock constructs, hs>$cls^{Y83X}$ and hs>$cls^{E189X}$, by site-directed mutagenesis from hs>$cls^{WT}$. The rationale behind creating the hs>$cls^{E189X}$ allele was that the mouse *Sox10del* and $Sox10^{Dom}$ mutations are almost in exactly the

some nucleotide position, but the mutant human protein lacks the additional novel amino acids created by the frameshift mutation in the mouse alleles. Furthermore, *Sox10del* was shown to encode a dominant negative protein in *in vitro* transcription assays (Potterf et al., 2000), whereas the $Sox10^{Dom}/^{+}$ phenotype can be fully explained by haploinsufficiency as was shown by comparison to a $Sox10^{lacZ}$/+ mouse (Britsch et al., 2001). To test whether any of these dominant phenotypes could be reproduced in zebrafish, in which lesions in *sox10* normally only cause a recessive phenotype, would represent an interesting future extension of this work.

The location of the lesion in the human Y83X allele is similar to our $cls^{t3}$ allele and truncates the protein upstream of the HMG domain. However, at the time this construct was made, the $cls^{t3}$ lesion had not been identified. The Y83X allele, together with E189X and K376X represented a convenient set of truncated proteins to investigate the functional abilities of Sox10 lacking each domain. Y83X and E189X truncate Sox10 approximately after the synergy domain and after the HMG domain respectively, whereas K376X only lacks the transactivation domain (Figure 7.11).

These constructs are prepared and ready to be assessed in overexpression studies and in a luciferase reporter construct assay as previously described for other heatshock constructs. Unfortunately, none of these experiments have been started to date.

Ectopic expression of either *sox9a* or *sox9b* was unable to rescue melanophores in *cls* mutants. It would be interesting to investigate rescue of other neural crest derivatives. The partial expressivity of the defect in sensory neurons of DRG compared to an essentially full expressivity of the pigment phenotype might indicate a functional redundancy with a similar protein in DRG. Very recently, K. Dutton has shown that injection of either *sox9a* or *sox9b* morpholino, which knocks-down *sox9a* and *sox9b* transcription, respectively, exacerbates the DRG defect in $cls^{t3}$ embryos (K. Dutton, unpubl. data). Thus, Sox9a and Sox9b might act redundantly with Sox10 within the DRG lineage.

Overexpression of Sox9a, Sox9b and mutant Sox10 proteins resulted in a higher proportion of deformed embryos. However, this result is likely to be artefactual, since the concentration of the various heatshock constructs was only estimated by gel electrophoresis and comparison to a quantitative DNA marker. Thus, some variation in the amount of construct injected could easily explain the differences in malformation observed.

# Mutagenesis primers to recreate human mutations

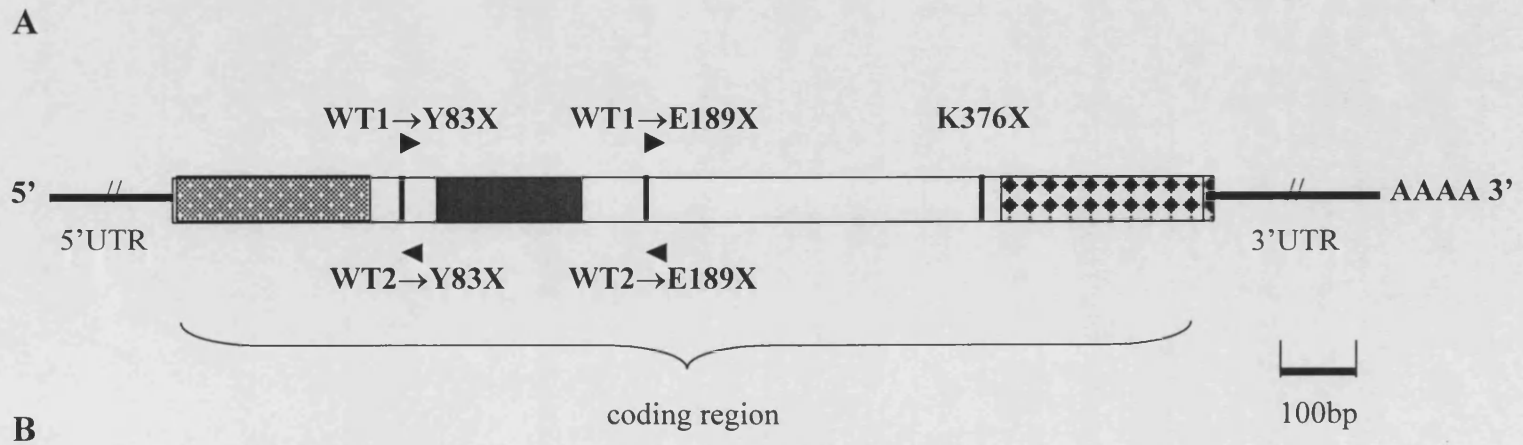


**Figure 7.11:** Location of mutagenesis primer pairs used to introduce mutant lesions identified in human alleles Y83X and E189X.
(A) The coding and untranslated regions (5' UTR, 5' untranslated region; 3' UTR, 3' untranslated region) of *sox10* are shown together with the HMG domain (black), the synergy domain (dotted) and the transactivation domain (checked). Long reverse complementary primers WT1→Y83X and WT2→Y83X were designed to introduce a premature stop codon truncating the protein at the C-terminal end of the synergy domain. Another primer pair WT1→E189X and WT2→E189X introduces a stop codon just downstream of the HMG domain and hence, the protein truncates after the synergy and HMG domain. The zebrafish mutant lesion found in $cls^{tw2}$ (K376X) is indicated. The scale bar represents the length of a 100bp fragment. (B) The diagram shows the 3 mutant heatshock constructs and their remaining functional domains. hs>$cls^{K376X}$ (hs>$cls^{tw2}$) only lacks the transactivation domain (checked), hs>$cls^{E189X}$ contains the entire N-terminal domain including the synergy (dotted) and the HMG domain (black), whereas hs>$cls^{Y83X}$ only contains the synergy domain (dotted). It is noted that these domains of zebrafish Sox10 are only putative and based on those defined in mouse Sox10.

In summary, we could demonstrate rescue of the melanophore lineage in *cls* mutants by ectopic expression of *sox10*, but not with mutant *cls* alleles, $cls^{m618}$ and $cls^{tw2}$, *sox9a* or *sox9b*. These results, together with the disrupted expression of *sox10* in *cls* mutants, the identification of 3 distinct *sox10* mutant lesions in *cls* alleles and tight linkage between the *cls* locus and *sox10* show beyond any reasonable doubt that *cls* encodes *sox10*. Hereafter, mutant alleles, for example $cls^{m618}$ and $cls^{tw2}$, will be renamed as $sox10^{m618}$ and $sox10^{tw2}$, respectively.

# Final discussion and conclusions

The aim of this project was to identify the gene encoded at the *colourless* locus; specifically, to test the candidate gene *sox10*, an HMG domain transcription factor. To address this objective we had to clone *sox10* and then test whether *cls* encoded *sox10*.

To clone *cls* we chose the candidate gene approach. Although there are several loci in mice and humans causing neurocristopathies with a similar combination of defects as observed in *cls*, detailed analysis of their phenotypes revealed *Sox10* as the most promising candidate gene. We chose RT-PCR and RACE PCR as the method to clone *sox10* using the then available mammalian *Sox10* sequences to design primers. The reliability of the primer design might have been improved, if the only recently published chicken *Sox10* could have been included. In the not too distant future, the entire zebrafish genome will have been sequenced and thus, cloning of zebrafish genes will be trivial.

Once a *sox10*-like cDNA was isolated, we were very cautious in assigning its true identity. *Sox* genes of the same subgroup show high similarities to each other and some *Sox* genes in the database were originally misidentified. Sequence alignments with other Sox10, Sox9 and Sox8 proteins, all members of subgroup E, and phylogenetic analysis both indicated that our clone represented a true Sox10 homologue. Furthermore, *in situ* hybridisation with a large fragment of this *sox10*-like gene also showed expression patterns highly reminiscent of other *Sox10* homologues in mouse, human and chicken. Lastly, our clone was placed next to the polymerase POLRF gene on LG3. *SOX10* on human chromosome 22q13.1 was very tightly linked to human POLRF consistent with our isolated cDNA being a *sox10* orthologue and with these genes defining a synteny between fish and human. Together, these data were in strong agreement with the proposition that the cloned cDNA represented a *sox10* homologue.

Next we presented very strong support in favour of the hypothesis that homozygous lesions in *sox10* caused the *cls* phenotype. First, we showed that *sox10* expression was disrupted in *cls* mutant embryos. Interestingly, *sox10* expression in all mutant alleles apart from $sox10^{ty22f}$ and $sox1010^{t3}$ was indistinguishable from wild-types up to approximately 24hpf. From 24hpf, *sox10* expression levels associated with cranial ganglia and in Schwann cells on the posterior lateral line nerve were reduced and *sox10*-

positive neural crest cells in the head and trunk showed a migrational defect. Thus, mutant *sox10* was first transcribed at normal levels, but the resulting mutant protein might have been unable to maintain its expression. Secondly, we demonstrated tight linkage between *sox10* and the *cls* mutant locus. Thirdly, we identified non-conservative point mutations in 3 of the *cls* alleles, which correlate with the mutant phenotype.

Molecular characterisation revealed that all three lesions were found in the coding region of *sox10*. L142Q in $sox10^{m618}$ affects a fully conserved residue in the HMG domain and is likely to diminish the ability of the mutant protein to bind to its DNA target sequences. K376X in $sox10^{tw2}$ and $sox10^{tw11}$ truncates the protein prematurely upstream of the transactivation domain. A fourth lesion was identified in $sox^{t3}$ by T. Carney, which caused truncation of the protein N-terminal to the HMG domain due to an insertion of a transposable element. No lesion has yet been found in $sox10^{ty22f}$, but it seems likely to be different to other previously identified alleles. $sox^{ty22f}$ transcripts were produced at lower levels at a stage when other alleles showed normal *sox10* expression. This indicates that the mutation might be in one of the regulatory elements or might destabilise the transcript.

The functional abilities of these 3 alleles were analysed in *in vitro* and *in vivo* studies. Consistent with our molecular characterisation of these mutant alleles, L142Q ($sox10^{m618}$) and K376X ($sox10^{tw2}$ and $sox10^{tw11}$) did not show any significant transcriptional activation ability in *in vitro* studies (W. Pavan, unpubl. data). Furthermore, unlike wild-type Sox10, mutant Sox10 proteins were unable to rescue the pigment phenotype *in vivo*. $sox10^{t3}$ was not included in such *in vivo* or *in vitro* experiments, since the lesion had not been identified at that time. Due to the very premature truncation of the protein, $sox10^{t3}$ is likely to be a functional null allele. Like $sox10^{t3}$, the human Y83X allele lacks the conserved DNA binding domain, both nuclear localisation signals and the transactivation domain, which leads to complete functional inactivation of the protein (Kuhlbrodt et al., 1998b).

Preliminary studies carried out by project students compared the severity of the *cls* mutant phenotype in various alleles and did not reveal any striking differences between any of the alleles. The phenotype of these alleles was also very similar to the maximal morphant phenocopy obtained with the highest doses of *sox10* morpholino (Dutton et al., 2001a). In strong *sox10* mutant phenotypes produced by the $sox10^{m618}$ or $sox10^{tw2/tw11}$ allele for example, melanophores with wild-type morphology are never present. However, even with the highest doses of *sox10* morpholino, an average of 11

normal melanophores were observed. Thus, taken together, these data suggest that *sox10*$^{m618}$, *sox10*$^{tw2}$ and *sox10*$^{tw11}$ are likely to be nulls (Dutton et al., 2001a). Consistent with being a null, mutant alleles failed to rescue pigment cells when ectopically expressed in *sox10* mutants despite the fact that under such circumstances, mutant proteins might be expected to be expressed at levels above endogenous expression levels. However, it remains to be tested whether mutant Sox10 protein can rescue other crest derivatives affected in *cls* mutants such as enteric ganglia or DRG. In accordance with the results from the *in vitro* studies testing transcriptional activation ability of mutant Sox10 proteins, failure to rescue pigment cells was most likely due to the inability to transactivate downstream targets such as *nacre*. In a similar experiment, the *Sox10*$^{Dom}$ mouse mutation and human *Sox10*$^{059}$ failed to induce endogenous *protein zero* (*Po*) expression in an *in vitro* cell line, a protein involved in regulating Schwann cell specific expression (Lemke et al., 1988; Peirano et al., 2000a).

The mammalian *Sox10* mutants showed a dominant phenotype, whilst only fully recessive phenotypes have been detected in *sox10* mutant alleles in zebrafish. Phenotypes in WS4 patients have been suggested to be generated via haploinsufficiency (Pingault et al., 1998a), but this has only been tested for the *Sox10*$^{Dom}$ mouse mutant. Heterozygous mutants in which the *Sox10* coding region was replaced with *lacZ* exhibited an identical phenotype to heterozygous *Sox10*$^{Dom/+}$ mice (Britsch et al., 2001). It is likely that the effects in these mice is due to the halving of the amount of Sox10 protein, thus suggesting the proposal that the *Sox10*$^{Dom/+}$ heterozygous phenotype is explained by the same mechanism. Heterozygous *cls* mutants however do not appear to have any defects. It will certainly be very interesting to investigate the cause of this discrepancy.

Dominant negativity has been suggested for several human alleles, but only conclusively shown *in vitro* for the artificially created *Sox10del* mouse allele (Potterf et al., 2000). None of the *cls* alleles appeared to be dominant negative when mutant protein was expressed ectopically. Likewise, uninjected heterozygous *cls* embryos also did not exhibit a lower number of pigment cells or show any kind of aberrant pigment cell morphology or distribution, which might be the result of dominant negatively acting Sox10 protein. However, the effect might be subtle and thus to assess dominant negativity conclusively, a concentration dependent effect of mutant Sox10 on the functional ability of co-transfected wild-type Sox10 would have to be demonstrated. For example, one could show wild-type Sox10 activating the *nacre* promoter in a luciferase

reporter construct. If any of the *cls* alleles were dominant negative, this activation by wild-type protein would be progressively inhibited by increasing amounts of co-transfected mutant protein. Alternatively, cultured undifferentiated neural crest cells might fail to survive or differentiate when transfected with mutant *sox10*.
The difference in dominance in mouse and fish might also be due to other factors such as genetic background. Thus, it will be interesting to investigate whether the artificially created mouse *Sox10del* allele exhibits any dominant negative function when overexpressed in zebrafish. *Sox10del* and the haploinsufficient $Sox10^{Dom}$ allele (Britsch et al., 2001), are caused by frameshift mutations which introduce 8 and 99 novel amino acids respectively, before premature truncation of the protein (Herbarth et al., 1998; Southard-Smith et al., 1998; Potterf et al., 2000). However, these extraneous residues are difficult to replicate and thus we have recreated the human E189X allele instead, which is truncated at almost exactly the same location as the two mouse alleles *Sox10del* (C190X) and $Sox10^{Dom}$ (E193insG; Pingault et al., 1998a; Southard-Smith et al., 1998; Potterf et al., 2000). No experiments with this artificial zebrafish *sox10* mutant allele have been carried out to date. However, this allele will hopefully be tested in our rescue assays for residual function and possible dominant negative effects by ectopic expression in wild-type embryos.

Finally, having identified *sox10*, we can now exploit the cell biological advantages of zebrafish to investigate the role of this transcription factor during neural crest development in detail. In neural crest cell transplant studies, it was shown that *sox10* acted cell autonomously in pigment cells (Kelsh and Eisen, 2000b), consistent with it being a transcription factor. Iontophoretic labelling of single neural crest cells revealed that ectomesenchymal fates differentiated normally, but non-ectomesenchymal crest derivatives died in *cls* embryos before differentiation. Furthermore, these cells were dying by an apoptotic mechanism within a discrete time window (Dutton et al., 2001b). Most neural crest cells in *cls* embryos showed restricted migration consistent with the counts of *sox10*-positive cells in *sox10 in situ* hybridisations described here.

Potentially, Sox10 could be involved in survival, migration, proliferation, differentiation or specification of neural crest cells. However, results from our *sox10 in situ* hybridisation studies suggest that the timing and distribution of *sox10* expression is consistent with a primary role for Sox10 in specification of neural crest derivatives.

A primary role for Sox10 in survival of neural crest cells seems unlikely since this would predict a very short time between the earliest defect and the onset of neural

crest cell death. At approximately 20hpf *cls* mutants are first distinguishable from wild-types by a lack of *nacre* expression, whereas apoptosis of neural crest cells in *cls* embryos is only observed approximately 15 hours later (Dutton et al., 2001b). In contrast, previous reports suggest that the delay between induction of apoptosis and detectable changes in cellular morphology is approximately 3-4 hours (Ikegami et al., 1999). Sox10 was shown to support survival of murine undifferentiated postmigratory neural cells *in vitro*. It was suggested that this function of Sox10 was mediated by regulating neuregulin (NRG1) signalling (Paratore et al., 2001). However, they failed to show that cell death due to lack of NRG1 signalling was the primary defect. Thus, this *in vitro* study is consistent with our observations and conclusions *in vivo* that survival is not Sox10's primary function.

A primary role of Sox10 in migration of neural crest cells is equally unlikely. Cranial neural crest cells are scattered across wild-type heads, whereas they tend to form clusters in *cls* mutants indicative of inhibited migration. Trunk neural crest cells fail to enter the lateral migration pathway and accumulate in the premigratory position. It could be argued that a migrational defect in crest derivatives lacking Sox10 might prevent them from encountering the appropriate environments for their trophic support and hence they die before differentiation. However, this does not hold true for the 3 pigment cell derivatives. The premigratory area is also a final location for those pigment cells which will form the dorsal stripe and thus, all the necessary factors have to be present. Despite this, fully differentiated pigment cells are not observed in the dorsal stripe of *cls* embryos.

Instead, we have proposed that Sox10's primary function is in specification of non-ectomesenchymal fates. In *cls* mutants the failure to become specified properly then leads to all other secondary defects observed such as inhibited migration, slightly reduced proliferation (K. Dutton and R. Kelsh, unpubl. data) and apoptosis prior to differentiation (Dutton et al., 2001b). This proposal fits well with observations of melanophore lineage markers in *cls* embryos (Dutton et al., 2001b; Figure 8.1). *Nacre*, like mammalian *Mitf* homologues, was shown to play an important role in specifying the melanophore fate and its expression was absent in *cls* embryos (Tachibana et al., 1996; Opdecamp et al., 1997; Lister et al., 1999). In fact, ectopic expression of *nacre* in *cls* embryos was sufficient to rescue melanophores (S. Elworthy, unpubl. data) consistent with the direct activation of the human *MITF* promoter by SOX10

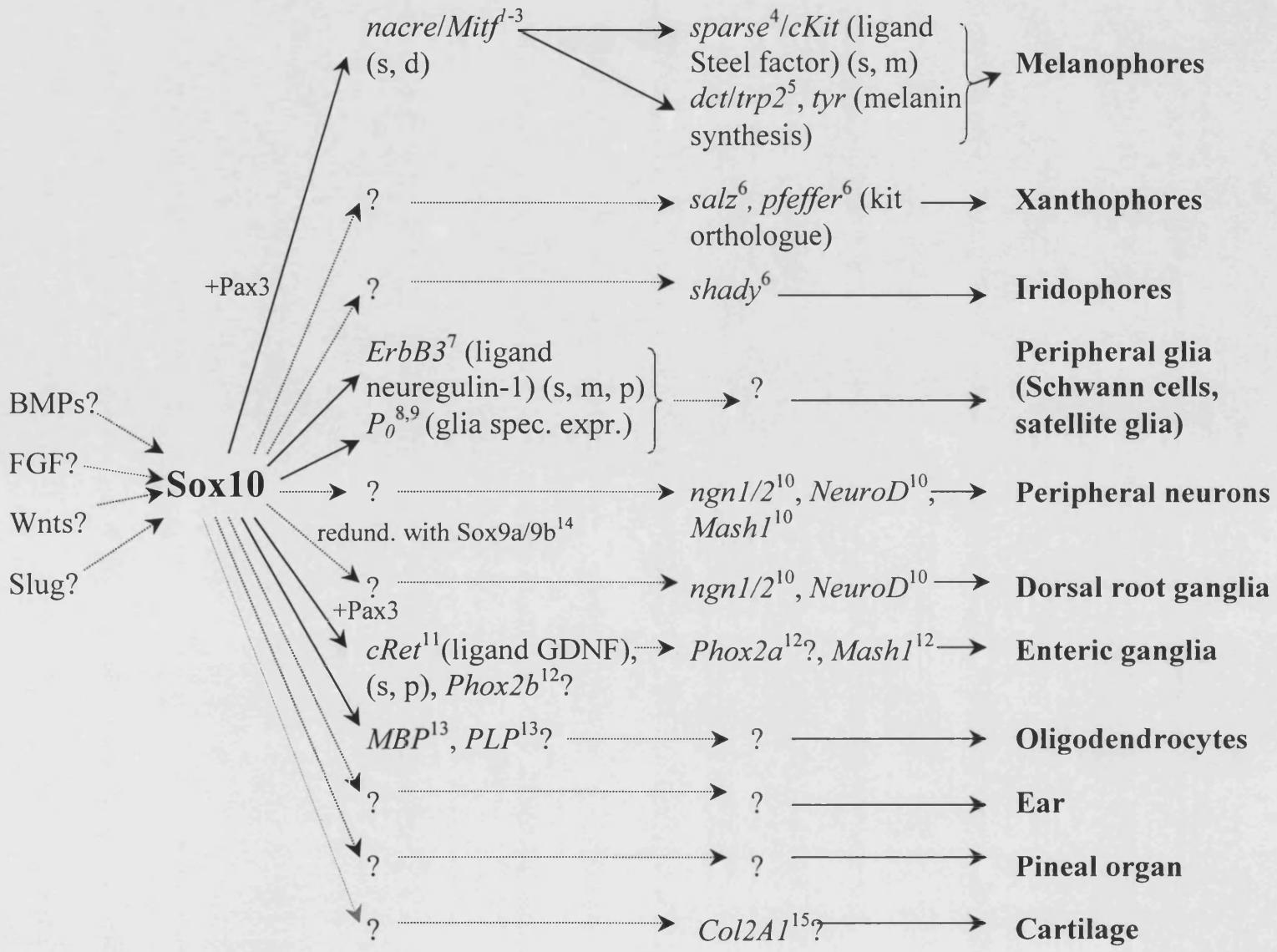

BMPs?
FGF?
Wnts?
Slug?
Sox10
+Pax3
+Pax3
redund. with Sox9a/9b[14]
nacre/Mitf[l-3] (s, d)
sparse[4]/cKit (ligand Steel factor) (s, m)
dct/trp2[5], tyr (melanin synthesis)
Melanophores
?
salz[6], pfeffer[6] (kit orthologue)
Xanthophores
?
shady[6]
Iridophores
ErbB3[7] (ligand neuregulin-1) (s, m, p)
P0[8,9] (glia spec. expr.)
?
Peripheral glia (Schwann cells, satellite glia)
?
ngn1/2[10], NeuroD[10], Mash1[10]
Peripheral neurons
?
ngn1/2[10], NeuroD[10]
Dorsal root ganglia
cRet[11](ligand GDNF), (s, p), Phox2b[12]?
Phox2a[12]?, Mash1[12]
Enteric ganglia
MBP[13], PLP[13]?
?
Oligodendrocytes
?
?
Ear
?
?
Pineal organ
?
Col2A1[15]?
Cartilage

**Figure 8.1:** Schematic of possible upstream and downstream targets of Sox10

Information on included interactions were taken from: [1] Lister et al., 1999, [2] Bondurand et al., 2000, [3] Potterf et al., 2000, [4] Parichy et al., 1999, [5] Kelsh et al., 2000c, [6] Kelsh et al., 1996, [7] Britsch et al., 2001, [8] Peirano et al., 2000a, [9] Peirano et al., 2000b, [10] Sommer et al., 1996, [11] Lang et al., 2000, [12] Pattyn et al., 1999, [13] Stolt et al., 2002, [14] K. Dutton, unpubl. observ., [15] Bell et al., 1997.

Solid arrows indicate a direct regulation, dashed arrows an unknown interaction. d, differentiation; m, migration; p, proliferation; s, survival; glia spec. expr., glia specific expression; redund., functionally redundant; BMP, Bone morphogenetic protein; dct, dopachrome tautomerase; FGF, fibroblast growth factor; GDNF, glial derived neurotrophic factor; MBP, myelin basic protein; Mitf, microphthalmia; ngn1/2, neurogenin 1/2; PLP, proteolipid protein; $P_0$, myelin protein zero; trp2, tyrosinase related protein-2; tyr, tyrosinase.

(Bondurand et al., 2000; Potterf et al., 2000). *Sparse*, a *c-kit* homologue, was shown to have a role in migration of melanoblasts and survival of both melanoblasts and melanophores and its expression was lost in *cls* and *nacre* embryos (Parichy et al., 1999; Kelsh et al., 2000c; Dutton et al., 2001b). The enzyme *Dopachrome tautomerase* (*Dct*; Kelsh et al., 2000c), an early melanoblast marker, is involved in synthesis of the melanin precursor and its expression is absent in *nacre* mutants (Lister et al., 1999). Furthermore, in mice, Sox10 was shown to activate the *Dct* promoter *in vitro* and was suggested to transiently regulate early *Dct* transcription directly (Potterf et al., 2001). Thus, the proposed model for Sox10 in melanophore specification included direct activation of *nacre*, likely in synergy with Pax3, which in turn directly or indirectly induced the expression of genes involved in migration, survival and differentiation of melanoblasts, such as *sparse* and *dct* (Dutton et al., 2001b).

Further tests of this model of Sox10 function will require the demonstration of similar scenarios to explain other neural crest defects, for example those in the enteric and sensory ganglia lineages. A schematic of suggested downstream targets of Sox10 in various neural crest derivatives is shown in Figure 8.1.

*cls* mutant embryos exhibit defects in only a subset of neural crest derivatives, the non-ectomesenchymal lineages, whereas ectomesenchymal derivatives appear unaffected (Kelsh and Eisen, 2000b; Dutton et al., 2001b). Our model suggests a role for Sox10 in specifying only the non-ectomesenchymal lineages. Thus, we asked whether *sox10* was expressed in all neural crest cells at a stage in which most cells are fate-restricted, or perhaps only in a subset of cells, which might represent non-ectomesenchymal precursors.

We first compared *sox10* expression to *crestin*, published as a pan-neural crest marker. However, at early stages we found that *crestin* was not expressed in cranial crest, which was subsequently confirmed by the original authors (P. Henion, pers. commun.). At the 16hpf stage, we observed a few scattered neural crest cells in the head starting to express *crestin*, but their fate restriction or specification is unknown. Nevertheless, *crestin* represents a useful marker for all trunk neural crest cells for future experiments.

Although *Sox10* has been used as a pan-neural crest marker in mouse (e.g. Pattyn et al., 1999), we showed in our double *in situ* labelling studies with *fkd6* and *sox10* that this was not true, at least in zebrafish. We observed some premigratory neural crest cells at the 6 somite and at the 14 somite stage, which were only labelled

with *fkd6*, but not *sox10,* demonstrating an incomplete overlap between these two crest markers.

One possible explanation for these *sox10*$^{-}$/*fkd6*$^{+}$ cells would be that they represent ectomesenchymal precursors. This hypothesis would be consistent with these fates being unaffected in *cls* mutant embryos (Kelsh and Eisen, 2000b). Craniofacial cartilage, an ectomesenchymal neural crest lineage, is only derived from the cranial neural crest, a region in which we had observed these *sox10*-negative cells. They appear to become distinct from non-ectomesenchymal lineages early during neural crest development (Schilling and Kimmel, 1994; Dutton et al., 2001b). Unfortunately, there is no marker for early ectomesenchymal precursors available to date. The hypothesis could then be confirmed by showing a non-overlapping signal with *sox10.*

In absence of a marker, one could hypothesise about the likely identity of these *sox10*$^{-}$ */fkd6*$^{+}$ cells depending on their location. At least 88% of cranial neural crest cells were shown to be fate-restricted and give rise to single cell-type fates (Schilling and Kimmel, 1994). Two research groups have investigated fate maps of cranial crest cell precursors, with slightly differing results. Schilling and colleagues found that the most lateral tier (1) of neural crest cells in a 6 somite stage embryo gave rise exclusively to neural cells, whereas Schwann cells, neural and pigment cells were derived from tier (2). More medial tiers (3-5) gave rise predominantly to pigment cells, glia and cartilage cells and clones from tier (6) were of an exclusively cartilage and connective tissue nature (Schilling and Kimmel, 1994). Although lateral cells never produced cartilage according to Schilling, Dorsky reported cartilage derived from both lateral and medial cells. Dorsky found that neurons were mostly derived from lateral cells, pigment from medial cells and glia from all positions (Dorsky et al., 1998). One possible explanation for this discrepancy is that Schilling only labelled the most superficial 20% of neural crest cells and thus, cartilage precursors observed by Dorsky could have been lateral cells in a deeper layer.

In our study in 6 somite stage embryos, *sox10*$^{-}$/*fkd6*$^{+}$ cranial neural crest cells could only be identified on the very surface due to visibility problems in deeper layers. Furthermore, the two cells documented in Figure 4.6A and B were located medially in tier (5) or (6) and thus could represent cartilage precursors according to (Schilling and Kimmel, 1994; Dorsky et al., 1998). This would support the hypothesis that *sox10* might not be expressed in ectomesenchymal derivatives and thus, they are unaffected in *sox10* mutant embryos.

Alternatively, the lack of defects in the cartilage lineage in *cls* mutants might be explained by functional redundancy of Sox10 with another protein in this cell-type. Functional redundancy might not only affect ectomesenchymal derivatives to prevent overt defects in *cls* mutants. It might also help to explain differences in severity of the DRG defect along the rostrocaudal axis or even the recessive phenotype observed in zebrafish *sox10* mutants in contrast to mammalian *Sox10* mutants. Such redundant action could be achieved between *sox10* and a second paralogue or *sox10* and a closely related *Sox* gene.

It is possible that a second paralogue of *sox10* exists in zebrafish that arose during the additional genome duplication event in ray-finned fish (Amores et al., 1998; Postlethwait et al., 1998). So far, for zebrafish *sox9* and *sox11*, such duplicated paralogues have been identified, but functional redundancy between those paralogues has not been investigated (Rimini et al., 1999; Chiang et al., 2001). However, functional redundancy between duplicated genes has been suggested for the two zebrafish *mitf* paralogues. It was shown that *mitfb* misexpression could rescue the formation of melanophores in *nacre* (*mitfa*) mutant embryos although at lower efficiency. Thus, Mitfb protein could functionally substitute for Mitfa (Lister et al., 2001).

There are several suitable approaches to investigate the existence of a second *sox10* paralogue. The degree of conservation between the pairs of Sox9 and Sox11 paralogues is high with 60% and 70% overall amino acid identity respectively (Rimini et al., 1999; Chiang et al., 2001). Thus, if they exist, the sequences of the two *sox10* paralogues would be expected to be very conserved at least within important functional domains. PCR-based methods could be designed that attempt to bias against the previously cloned *sox10* and thus, favour amplification of a second paralogue. Firstly, one could perform a PCR amplification on $sox10^{t3}$ cDNA with degenerate primers designed against sequences conserved in *sox10*, but not *sox9* regions. *sox10* expression in the $sox10^{t3}$ allele is drastically reduced (chapter 6) and thus contains less *sox10a* transcript that could serve as a template. Secondly, a PCR amplification could be performed on $sox10^{t3}$ cDNA with degenerate *sox10*-specific primers designed to flank the t3 insertion site. PCR products derived from the previously cloned *sox10a* would be distinguished from *sox10b* derived products by an 1.4kb increase in size, corresponding to the insertion of the transposable element. Alternatively, a *sox10b* fragment could be amplified using degenerate *sox10*-specific primers and genomic DNA from the c1033

line. The latter is a deletion mutant, which lacks markers surrounding *sox10a* and thus is likely also to lack *sox10a* itself.

Another consideration when trying to identify a putative *sox10b* is the time of its expression. To successfully isolate *sox10b*, one has to guess when it might be expressed during development. To use one of the first two approaches, cDNA from *sox10*$^{t3}$ embryos would have to be isolated from various timepoints. However, cDNA from *cls* mutant adults cannot be obtained, since *cls* alleles are embryonic lethal. The third approach, PCR from genomic DNA, avoids this problem and might thus be the most promising method.

Once cloned, one would have to test whether this putative *sox10b* was expressed at the correct stage to be able to act redundantly with Sox10a. Next, a functional redundancy between Sox10a and Sox10b could be investigated by ectopic *sox10b* expression in *cls* embryos. If they acted redundantly, one might expect to see partial rescue of the *cls* phenotype. The demonstration of the enhancement of the *cls* phenotype by injection of a *sox10b* morpholino into *cls* embryos would also suggest functional redundancy between the Sox10 paralogues.

Even in the absence of a *sox10b*, a closely related *Sox* gene might act functionally redundantly with *sox10*. Such an action in the ectomesenchymal lineage might explain the lack of defects in these cell-types in *cls* mutants. In addition or alternatively, this close relative might substitute for lack of *sox10* in DRG development. In *cls* embryos, DRGs in anterior segments are much less severely affected than in posterior segments (Kelsh and Eisen, 2000b) which might suggest a possible redundancy.

*Sox* genes of subgroup E (*sox8*, sox9 and *sox10* and likely any possible duplicated paralogues) are all closely related on the basis of their sequence and gene structure (Bowles et al., 2000). The regions of their expression are overlapping in the neural crest (*sox8*, *sox9* and *sox10*), in the otic vesicle (*sox9*, *sox10*), in the eye (*sox8*, *sox9*), in DRGs (*sox8*, *sox10*), limbs or pectoral fin buds (*sox8*, *sox9*, *sox10*), somites (*sox8*, *sox9*), CNS (*sox8*, *sox9*, *sox10*), gut (*sox8*, *sox9*, *sox10*), testis (*sox8*, *sox9*) and the heart (*sox8*, *sox9*, *sox10*) (Wright et al., 1995; Ng et al., 1997; Bondurand et al., 1998b; Herbarth et al., 1998; Kuhlbrodt et al., 1998a; Southard-Smith et al., 1998; Bell et al., 2000; Cheng et al., 2000; Schepers et al., 2000; Chiang et al., 2001; Dutton et al., 2001b). Together, these facts could suggest some functional redundancy between these proteins, although this has not been demonstrated between *Sox* genes in any organism.

In chicken, closely related Sox1, Sox2 and Sox3 proteins, which are involved in lens development, were all shown to bind to and activate the DC5 minimal enhancer of the chicken δ1-*crystallin* gene (Kamachi et al., 1995). This makes functional redundancy between them likely, although no such studies have yet been carried out to test this. If Sox10 functioned partially redundantly with either Sox9a or Sox9b we would predict overlapping expression domains.

To begin to test this, we performed double *in situ* hybridisations with *sox10* and *sox9b* which revealed a pattern of incomplete overlap of expression. *sox9b*$^{+}$/*sox10*$^{-}$ cells tended to be found in lateral cranial neural crest regions in 6 somite stage embryos, a region thought to give rise to mostly neural derivatives (Schilling and Kimmel, 1994). At the 14 somite stage, the majority of migratory cranial and premigratory trunk neural crest cells were double labelled, although some cells above the brain and anterior to the otic placode clearly expressed only *sox10* or *sox9b*, similarly to double labelling with *fkd6* and *sox10*. It is unclear, how this differential labelling could correspond to a distinct specification status of each cell.

In chapter 7 we have described the failure to rescue the melanophore phenotype in *cls* mutants by ectopic *sox9b* expression, which represents evidence against functional redundancy at least in pigment precursors. However, from our *in situ* hybridisation results, a redundancy in neuronal derivatives for example in DRGs or in non-ectomesenchymal derivatives in general, seems possible.

Alternatively, Sox10 could exhibit a repressive function on Sox9b in specifying ectomesenchymal fates. Thus, double-labelled cells would contribute to non-ectomesenchymal fates, and the few *sox9b*$^{+}$/*sox10*$^{-}$ cells represented cartilage precursors. In *cls* mutants, these latter fates indeed form normally, whereas non-ectomesenchymal derivatives fail to be specified properly. If this hypothesis held true, one would expect to be able to demonstrate rescue of non-ectomesenchymal crest precursors with ectopic *sox10* alone and when co-injected with *sox9b*. If *sox9b* alone was overexpressed in *cls* mutants, an increase of ectomesenchymal derivatives might be expected to be seen. However, this has not yet been investigated to date.

In double *in situ* hybridisations with *sox10* and *sox9a*, expression domains did not seem to broadly overlap, but rather appeared complementary. It was difficult to determine whether signals were complementary or overlapping in adjacent regions, because the expression signals for both probes were very strong. To confirm non-overlap or complementarity of expression domains, future experiments should include

sectioning of these embryos. We did not investigate the developmental potential of these *sox9a*-positive cells, but they were found in a position approximately where the otic placode will be formed 3 hours later. In 14 somite stage embryos, it was impossible to determine whether *sox9a* was expressed weakly in any neural crest cells. However, if *sox9a* was expressed at reasonable levels, it should have been possible to observe some purple staining as was seen in the equivalent *sox9b/sox10* double *in situ* hybridisations (compare Figure 4.8G to Figure 4.9G). Likewise, the precise expression domains for *sox9a* in the otic placode were difficult to establish without sectioning the specimen. Overall, it appeared that *sox9a* and *sox10* expression domains, in contrast to *sox9b* and *sox10*, did not overlap broadly although further tests are required to confirm this tentative result.

Only recently, K. Dutton was able to demonstrate a redundant action of Sox10 and Sox9a and Sox9b in DRG. *sox9a* and *sox9b* morpholino injections into *cls* mutants each exacerbated the DRG phenotype of *cls* mutants indicating functional redundancy between Sox9a, Sox9b and Sox10 in this cell lineage (K. Dutton, unpubl. data). On the other hand, neither *sox9a* nor *sox9b* was able to rescue melanophores. Thus, this functional redundancy appears to be specific to the DRG cell lineage and does not extend to the pigment derivatives. Elucidating the exact role and interplay of Sox10 with each of the Sox9 paralogues in different cell lineages will be an interesting and rewarding line of future research.

Overall, the identification of *colourless* as a zebrafish *sox10* homologue reinforces the value of this fish model for human Waardenburg Shah Syndrome, since neural crest development in zebrafish and other vertebrates shares many characteristics (Raible et al., 1992). The work described herein opened up the possibility of utilizing the cell biological advantages of this model organism to elucidate the role of Sox10 during neural crest development *in vivo* at a level of resolution difficult to achieve in other model species and to look at evolutionary conservation of Sox10 function.

# Bibliography


**Albers, B.** (1987). Competence as the main factor determining the size of the neural plate. *Dev. Growth Diff.* **29**: 535-545.

**Ambrosetti, D. C., C. Basilico and L. Dailey** (1997). Synergistic activation of the *fibroblast growth factor 4* enhancer by Sox2 and Oct-3 depends on protein-protein interactions facilitated by a specific spatial arrangement of factor binding sites. *Mol. Cell. Biol.* **17**: 6321-6329.

**Amores, A., A. Force, Y. L. Yan, L. Joly, C. Amemiya, A. Fritz, R. K. Ho, J. Langeland, V. Prince, Y. L. Wang, M. Westerfield, M. Ekker and J. H. Postlethwait** (1998). Zebrafish *hox* clusters and vertebrate genome evolution. *Science* **282**: 1711-1714.

**Amores, A. and J. H. Postlethwait** (1999). Banded Chromosomes and the Zebrafish Karyotype. *Meth. Cell Biol.*, Academic Press. **60:** 323-338.

**Appel, B. and J. S. Eisen** (1998). Regulation of neuronal specification in the zebrafish spinal cord by Delta function. *Development* **125**: 371-380.

**Asher, J. H., Jr., A. Sommer, R. Morell and T. B. Friedman** (1996). Missense mutation in the paired domain of *PAX3* causes craniofacial-deafness-hand syndrome. *Hum. Mutat.* **7**: 30-35.

**Auerbach, R.** (1954). Analysis of the developmental effects of a lethal mutation in the house mouse. *J. Exp. Zool.* **127**: 305-330.

**Baker, C., M. Bronner-Fraser, N. Le Douarin and M. A. Teillet** (1997a). Early- and late-migrating cranial neural crest cell populations have equivalent developmental potential *in vivo*. *Dev. Biol.* **186**: B72.

**Baker, C. V. and M. Bronner-Fraser** (1997b). The origins of the neural crest. Part I: embryonic induction. *Mech. Dev.* **69**: 3-11.

**Baroffio, A., E. Dupin and N. M. Le Douarin** (1991). Common precursors for neural and mesectodermal derivatives in the cephalic neural crest. *Development* **112**: 301-305.

**Barth, K. A., Y. Kishimoto, K. B. Rohr, C. Seydler, S. Schulte-Merker and S. W. Wilson** (1999). Bmp activity establishes a gradient of positional information throughout the entire neural plate. *Development* **126**: 4977-4987.

**Barton, D. E., B. S. Kwon and U. Francke** (1988). Human *tyrosinase* gene, mapped to chromosome 11 (q14-q21), defines second region of homology with mouse chromosome 7. *Genomics* **3**: 17-24.

**Baynash, A. G., K. Hosoda, A. Giaid, J. A. Richardson, N. Emoto, R. E. Hammer and M. Yanagisawa** (1994). Interaction of endothelin-3 with endothelin-B receptor is essential for development of epidermal melanocytes and enteric neurons. *Cell* **79**: 1277-1285.

**Bell, D. M., K. K. H. Leung, S. C. Wheatley, L. J. Ng, S. Zhou, K. W. Ling, M. H. Sham, P. Koopman, P. P. L. Tam and K. S. E. Cheah** (1997). SOX9 directly regulates the type-II collagen gene. *Nat. Genet.* **16**: 174-178.

**Bell, K. M., P. S. Western and A. H. Sinclair** (2000). *SOX8* expression during chick embryogenesis. *Mech. Dev.* **94**: 257-260.

**Bianchi, M. E., M. Beltrame and G. Paonessa** (1989). Specific recognition of cruciform DNA by nuclear protein HMG1. *Science* **243**: 1056-1059.

**Bisgrove, B. W., D. W. Raible, V. Walter, J. S. Eisen and D. J. Grunwald** (1997). Expression of *c-ret* in the zebrafish embryo: Potential roles in motoneuronal development. *J. Neurobiol.* **33**: 749-768.

**Boissy, R. E. and J. J. Nordlund** (1997). Molecular Basis of Congenital Hypopigmentary Disorders in Humans: A Review. *Pig. Cell Res.* **10**: 12-24.

**Boissy, R. E., H. Zhao, W. S. Oetting, L. M. Austin, S. C. Wildenberg, Y. L. Boissy, Y. Zhao, R. A. Sturm, V. J. Hearing, R. A. King and J. J. Nordlund** (1996). Mutation in and lack of expression of *tyrosinase-related protein-1* (*TRP- 1*) in melanocytes from an individual with brown oculocutaneous albinism: a new subtype of albinism classified as "OCA3". *Am. J. Hum. Genet.* **58**: 1145-1156.

**Bondurand, N., A. Kobetz, V. Pingault, N. Lemort, F. Encha-Razavi, G. Couly, D. E. Goerich, M. Wegner, M. Abitbol and M. Goossens** (1998a). Expression of the *SOX10* gene during human development. *FEBS Lett.* **432**: 168-172.

**Bondurand, N., K. Kuhlbrodt, V. Pingault, J. Enderich, M. Sajus, N. Tommerup, M. Warburg, R. C. M. Hennekam, A. P. Read, M. Wegner and M. Goossens** (1999). A molecular analysis of the Yemenite deaf-blind hypopigmentation syndrome: SOX10 dysfunction causes different neurocristopathies. *Hum. Mol. Genet.* **8**: 1785-1789.

**Bondurand, N., V. Pingault, D. E. Goerich, N. Lemort, E. Sock, C. L. Caignec, M. Wegner and M. Goossens** (2000). Interaction among *SOX10*, *PAX3* and *MITF*, three genes altered in Waardenburg syndrome. *Hum. Mol. Genet.* **9**: 1907-1917.

**Bondurand, N., V. Pingault, A. Kobetz, N. Lemort, M. O. Prehu, M. Sajus, D. E. Goerich, M. Wegner, M. Abitbol and M. Goossens** (1998b). Structure and expression of human SOX10, a transcriptional regulator involved in neural crest development. *Europ. J. Hum. Genet.* **6**: 4113.

**Bowles, J., G. Schepers and P. Koopman** (2000). Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. *Dev. Biol.* **227**: 239-255.

**Britsch, S., D. E. Goerich, D. Riethmacher, R. I. Peirano, M. Rossner, K. A. Nave, C. Birchmeier and M. Wegner** (2001). The transcription factor Sox10 is a key regulator of peripheral glial development. *Gen. Dev.* **15**: 66-78.

**Bronner-Fraser, M.** (1986). Analysis of the early stages of trunk neural crest migration in avian embryos using monoclonal antibody HNK-1. *Dev. Biol.* **115**: 44-55.

**Burns, A. J. and N. M. Le Douarin** (1998). The sacral neural crest contributes neurons and glia to the post-umbilical gut: spatiotemporal analysis of the development of the enteric nervous system. *Development* **125**: 4335-4347.

**Chakravarti, A.** (1996). Endothelin receptor-mediated signaling in Hirschsprung disease. *Hum. Mol. Genet.* **5**: 303-307.

**Chang, C. B. and A. Hemmati-Brivanlou** (1998). Neural crest induction by Xwnt7B in *Xenopus*. *Dev. Biol.* **194**: 129-134.

**Cheng, Y., M. Cheung, M. M. Abu-Elmagd, A. Orme and P. J. Scotting** (2000). Chick *Sox10*, a transcription factor expressed in both early neural crest cells and central nervous system. *Brain Res. Dev. Brain Res.* **121**: 233-241.

**Chiang, E. F., C. I. Pai, M. Wyatt, Y. L. Yan, J. Postlethwait and B. Chung** (2001). Two *sox9* genes on duplicated zebrafish chromosomes: expression of similar transcription activators in distinct sites. *Dev. Biol.* **231**: 149-163.

**Cornell, R. A. and J. S. Eisen** (2000). Delta signaling mediates segregation of neural crest and spinal sensory neurons from zebrafish lateral neural plate. *Development* **127**: 2873-2882.

**Dailey, L., H. Yuan and C. Basilico** (1994). Interaction between a novel F9-specific factor and octamer-binding proteins is required for cell-type-restricted activity of the fibroblast growth factor 4 enhancer. *Mol. Cell. Biol.* **14**: 7758-7769.

**Deol, M. S., G. M. Truslove and A. McLaren** (1986). Genetic activity at the *albino* locus in Cattanach's insertion in the mouse. *J. Embryol. Exp. Morphol.* **96**: 295-302.

**Dorsky, R. I., R. T. Moon and D. W. Raible** (1998). Control of neural crest cell fate by the Wnt signalling pathway. *Nature* **396**: 370-373.

**Driever, W., L. Solnica-Krezel, A. F. Schier, S. C. Neuhauss, J. Malicki, D. L. Stemple, D. Y. Stainier, F. Zwartkruis, S. Abdelilah, Z. Rangini, J. Belak and C. Boggs** (1996). A genetic screen for mutations affecting embryogenesis in zebrafish. *Development* **123**: 37-46.

**Durbec, P. L., L. B. Larsson-Blomberg, A. Schuchardt, F. Costantini and V. Pachnis** (1996). Common origin and developmental dependence on *c-ret* of subsets of enteric and sympathetic neuroblasts. *Development* **122**: 349-358.

**Dutton, K., J. R. Dutton, A. Pauliny and R. N. Kelsh** (2001a). A morpholino phenocopy of the *colourless* mutant. *Genesis* **30**: 188-189.

**Dutton, K. A., A. Pauliny, S. S. Lopes, S. Elworthy, T. J. Carney, J. Rauch, R. Geisler, P. Haffter and R. N. Kelsh** (2001b). Zebrafish *colourless* encodes *sox10* and specifies non-ectomesenchymal neural crest fates. *Development* **128**: 4113-4125.

**Edery, P., T. Attie, J. Amiel, A. Pelet, C. Eng, R. M. Hofstra, H. Martelli, C. Bidaud, A. Munnich and S. Lyonnet** (1996). Mutation of the *endothelin-3* gene in the Waardenburg-Hirschsprung disease (Shah-Waardenburg syndrome). *Nat. Genet.* **12**: 442-444.

**Edery, P., S. Lyonnet, L. M. Mulligan, A. Pelet, E. Dow, L. Abel, S. Holder, C. Nihoul-Fekete, B. A. Ponder and A. Munnich** (1994). Mutations of the *RET* proto-oncogene in Hirschsprung's disease. *Nature* **367**: 378-380.

**Erickson, C. A., T. D. Duong and K. W. Tosney** (1992). Descriptive and experimental analysis of the dispersion of neural crest cells along the dorsolateral path and their entry into ectoderm in the chick embryo. *Dev. Biol.* **151**: 251-272.

**Ferrari, S., V. R. Harley, A. Pontiggia, P. N. Goodfellow, R. Lovell-Badge and M. E. Bianchi** (1992). SRY, like HMG1, recognizes sharp angles in DNA. *Embo J.* **11**: 4497-4506.

**Force, A., M. Lynch, F. B. Pickett, A. Amores, Y. L. Yan and J. Postlethwait** (1999). Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**: 1531-1545.

**Ford, G.** (2000). Quantitative analysis of neural crest defects in *colourless* mutant embryos. Bath, Department of Biology and Biochemistry, University of Bath.

**Foster, J. W., M. A. Dominguez-Steglich, S. Guioli, G. Kowk, P. A. Weller, M. Stevanovic, J. Weissenbach, S. Mansour, I. D. Young and P. N. Goodfellow** (1994). Campomelic dysplasia and autosomal sex reversal caused by mutations in an *SRY*-related gene. *Nature* **372**: 525-530.

**Gardner, J. M., Y. Nakatsu, Y. Gondo, S. Lee, M. F. Lyon, R. A. King and M. H. Brilliant** (1992). The mouse *pink-eyed dilution* gene: association with human Prader-Willi and Angelman syndromes. *Science* **257**: 1121-1124.

**Gariepy, C. E., S. C. Williams, J. A. Richardson, R. E. Hammer and M. Yanagisawa** (1998). Transgenic expression of the endothelin-B receptor prevents congenital intestinal aganglionosis in a rat model of Hirschsprung disease. *J. Clin. Invest.* **102**: 1092-1101.

**Giebel, L. B. and R. A. Spritz** (1991). Mutation of the *KIT* (mast/stem cell growth factor receptor) proto-oncogene in human piebaldism. *Proc. Natl. Acad. Sci. USA* **88**: 8696-8699.

**Girard, F., F. Cremazy, P. Berta and A. Renucci** (2001). Expression pattern of the *Sox31* gene during zebrafish embryonic development. *Mech. Dev.* **100**: 71-73.

**Goulding, M., S. Sterrer, J. Fleming, R. Balling, J. Nadeau, K. J. Moore, S. D. Brown, K. P. Steel and P. Gruss** (1993). Analysis of the *Pax-3* gene in the mouse mutant splotch. *Genomics* **17**: 355-363.

**Goulding, M. D., G. Chalepakis, U. Deutsch, J. R. Erselius and P. Gruss** (1991). Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *Embo J.* **10**: 1135-1147.

**Gross, A., J. Kunze, R. F. Maier, G. Stoltenburg-Didinger, I. Grimmer and R. Obladen** (1995). Autosomal-recessive neural crest syndrome with Albinism, black lock, cell-migration disorder of the neurocytes of the gut, and deafness-ABCD Syndrome. *Am. J. Med. Genet.* **56**: 322-326.

**Grosschedl, R., K. Giese and J. Pagel** (1994). HMG domain proteins: architectural elements in the assembly of nucleoprotein structures. *Trends Genet.* **10**: 94-100.

**Haffter, P., M. Granato, M. Brand, M. C. Mullins, M. Hammerschmidt, D. A. Kane, J. Odenthal, F. J. M. vanEeden, Y. J. Jiang, C. P. Heisenberg, R. N. Kelsh, M. Furutani-Seiki, E. Vogelsang, D. Beuchle, U. Schach, C. Fabian and C. Nuesslein-Volhard** (1996a). The identification of genes with unique and essential functions in the development of the zebrafish, *Danio rerio*. *Development* **123**: 1-36.

**Haffter, P., J. Odenthal, M. C. Mullins, S. Lin, M. J. Farrell, E. Vogelsang, F. Haas, M. Brand, F. J. M. vanEeden, M. Furutani-Seiki, M. Granato, M. Hammerschmidt, C. P. Heisenberg, Y. J. Jiang, D. A. Kane, R. N. Kelsh, N. Hopkins and C. Nuesslein-Volhard** (1996b). Mutations affecting pigmentation and shape of the adult zebrafish. *Dev. Gen. Evol.* **206**: 260-276.

**Halloran, M. C., M. Sato-Maeda, J. T. Warren, F. Su, Z. Lele, P. H. Krone, J. Y. Kuwada and W. Shoji** (2000). Laser-induced gene expression in specific cells of transgenic zebrafish. *Development* **127**: 1953-1960.

**Hammerschmidt, M., P. Blader and U. Straehle** (1999). Strategies to perturb zebrafish development. *Meth. Cell Biol.*, Academic Press. **59:** 87-115.

**Harley, V. R., R. Lovell-Badge and P. N. Goodfellow** (1994). Definition of a consensus DNA binding site for SRY. *Nucleic Acids Res* **22**: 1500-1501.

**Hemesath, T. J., E. Steingrimsson, G. McGill, M. J. Hansen, J. Vaught, C. A. Hodgkinson, H. Arnheiter, N. G. Copeland, N. A. Jenkins and D. E. Fisher** (1994). *microphthalmia*, a critical factor in melanocyte development, defines a discrete transcription factor family. *Gen. Dev.* **8**: 2770-2780.

**Henion, P. D. and J. A. Weston** (1997). Timing and pattern of cell fate restrictions in the neural crest lineage. *Development* **124**: 4351-4359.

**Herbarth, B., V. Pingault, N. Bondurand, K. Kuhlbrodt, I. Hermans-Borgmeyer, A. Puliti, N. Lemort, M. Goossens and M. Wegner** (1998). Mutation of the *Sry*-related *Sox10* gene in *Dominant megacolon*, a mouse model for human Hirschsprung disease. *Proc. Natl. Acad. Sci. USA* **95**: 5161-5165.

**Hirschsprung, H.** (1888). Stuhltraegheit Neugeborener in Folge von Dilatation und Hypertrophie des Colons. *Jahrb. Kinderheilk.* **27**: 1-7.

**His, W.** (1868). *Untersuchungen ueber die erste Anlage des Wirbeltierleibes. Die erste Entwicklung des Huehnchens im Ei.* Leipzig, F.C.W. Vogel.

**Hodgkinson, C. A., K. J. Moore, A. Nakayama, E. Steingrimsson, N. G. Copeland, N. A. Jenkins and H. Arnheiter** (1993). Mutations at the mouse *microphthalmia* locus are associated with defects in a gene encoding a novel basic-helix-loop-helix-zipper protein. *Cell* **74**: 395-404.

**Hofstra, R. M., J. Osinga, G. Tan-Sindhunata, Y. Wu, E. J. Kamsteeg, R. P. Stulp, C. van Ravenswaaij-Arts, D. Majoor-Krakauer, M. Angrist, A. Chakravarti, C. Meijers and C. H. Buys** (1996). A homozygous mutation in the *endothelin-3* gene associated with a combined Waardenburg type 2 and Hirschsprung phenotype (Shah- Waardenburg syndrome). *Nat. Genet.* **12**: 445-447.

**Hornyak, T. J., D. J. Hayes, L. Chiu and E. B. Ziff** (2001). Transcription factors in melanocyte development: distinct roles for Pax- 3 and Mitf. *Mech. Dev.* **101**: 47-59.

**Hosoda, K., R. E. Hammer, J. A. Richardson, A. G. Baynash, J. C. Cheung, A. Giaid and M. Yanagisawa** (1994). Targeted and natural (*piebald-lethal*) mutations of *endothelin-B receptor* gene produce megacolon associated with spotted coat color in mice. *Cell* **79**: 1267-1276.

**Hoth, C. F., A. Milunsky, N. Lipsky, R. Sheffer, S. K. Clarren and C. T. Baldwin** (1993). Mutations in the paired domain of the human *PAX3* gene cause Klein-Waardenburg syndrome (WS-III) as well as Waardenburg syndrome type I (WS-I). *Am. J. Hum. Genet.* **52**: 455-462.

**Hukriede, N. A., L. Joly, M. Tsang, J. Miles, P. Tellis, J. A. Epstein, W. B. Barbazuk, F. N. Li, B. Paw, J. H. Postlethwait, T. J. Hudson, L. I. Zon, J. D. McPherson, M. Chevrette, I. B. Dawid, S. L. Johnson and M. Ekker** (1999). Radiation hybrid mapping of the zebrafish genome. *Proc. Natl. Acad. Sci. USA* **96**: 9745-9750.

**Hulten, M. A., M. M. Honeyman, A. J. Mayne and M. J. Tarlow** (1987). Homozygosity in *piebald* trait. *J Med Genet* **24**: 568-571.

**Ikegami, R., P. Hunter and T. D. Yager** (1999). Developmental activation of the capability to undergo checkpoint- induced apoptosis in the early zebrafish embryo. *Dev. Biol.* **209**: 409-433.

**Inoue, K., Y. Tanabe, J. Wilson, S. Hirabayashi, M. Wegner and J. R. Lupski** (1999). Myelin deficiencies in both the central and peripheral nervous systems associated with *SOX10* transcription factor mutations. *Am. J. Hum. Genet.* **65**: 13.

**Jantzen, H. M., A. Admon, S. P. Bell and R. Tjian** (1990). Nucleolar transcription factor hUBF contains a DNA-binding motif with homology to HMG proteins. *Nature* **344**: 830-836.

**Joyner, A. L. and G. R. Martin** (1987). *En-1* and *En-2*, two mouse genes with sequence homology to the *Drosophila engrailed* gene: expression during embryogenesis. *Gen. Dev.* **1**: 29-38.

**Kadesch, T.** (1993). Consequences of heteromeric interactions among helix-loop-helix proteins. *Cell. Growth Differ.* **4**: 49-55.

**Kamachi, Y., K. S. E. Cheah and H. Kondoh** (1999). Mechanism of regulatory target selection by the SOX high-mobility-group domain proteins as revealed by comparison of SOX1/2/3 and SOX9. *Mol. Cell. Biol.* **19**: 107-120.

**Kamachi, Y., S. Sockanathan, Q. Liu, M. Breitman, R. Lovell-Badge and H. Kondoh** (1995). Involvement of SOX proteins in lens-specific activation of crystallin genes. *Embo J.* **14**: 3510-3519.

**Kamachi, Y., M. Uchikawa and H. Kondoh** (2000). Pairing SOX off with partners in the regulation of embryonic development. *Trends Genet.* **16**: 182-187.

**Kapur, R. P., R. Livingston, B. Doggett, D. A. Sweetser, J. R. Siebert and R. D. Palmiter** (1996). Abnormal microenvironmental signals underlie intestinal aganglionosis in *Dominant megacolon* mutant mice. *Dev. Biol.* **174**: 360-369.

**Kapur, R. P., D. A. Sweetser, B. Doggett, J. R. Siebert and R. D. Palmiter** (1995). Intercellular signals downstream of Endothelin receptor-B mediate colonization of the large intestine by enteric neuroblasts. *Development* **121**: 3787-3795.

**Kapur, R. P., C. Yost and R. D. Palmiter** (1992). A transgenic model for studying development of the enteric nervous system in normal and aganglionic mice. *Development* **116**: 167-175.

**Kapur, R. P., C. Yost and R. D. Palmiter** (1993). Aggregation chimeras demonstrate that the primary defect responsible for aganglionic megacolon in *lethal spotted* mice is not neuroblast autonomous. *Development* **117**: 993-999.

**Kelsh, R. N., M. Brand, Y. J. Jiang, C. P. Heisenberg, S. Lin, P. Haffter, J. Odenthal, M. C. Mullins, F. J. M. vanEeden, M. Furutani-Seiki, M. Granato, M. Hammerschmidt, D. A. Kane, R. M. Warga, D. Beuchle, L. Vogelsang and C. Nuesslein-Volhard** (1996). Zebrafish pigmentation mutations and the processes of neural crest development. *Development* **123**: 369-389.

**Kelsh, R. N., K. Dutton, J. Medlin and J. S. Eisen** (2000a). Expression of zebrafish *fkd6* in neural crest-derived glia. *Mech. Dev.* **93**: 161-164.

**Kelsh, R. N. and J. S. Eisen** (2000b). The zebrafish *colourless* gene regulates development of non-ectomesenchymal neural crest derivatives. *Development* **127**: 515-525.

**Kelsh, R. N., B. Schmid and J. S. Eisen** (2000c). Genetic analysis of melanophore development in zebrafish embryos. *Dev. Biol.* **225**: 277-293.

**Kemphues, K. J., J. R. Priess, D. G. Morton and N. S. Cheng** (1988). Identification of genes required for cytoplasmic localization in early *C. elegans* embryos. *Cell* **52**: 311-320.

**Kent, J.** (1996). A male specific role for Sox9 in vertebrate sex determination. *Development* **122**: 2813-2822.

**Kimmel, C. B., W. W. Ballard, S. R. Kimmel, B. Ullmann and T. F. Schilling** (1995). Stages of Embryonic Development of the Zebrafish. *Dev. Dyn.* **203**: 253-310.

**King, R. A., V. J. Hearing, D. J. Creel and W. S. Oetting** (1995). Albinism. *The Metabolic and Molecular Bases of Inherited Disease*. C. R. Scriver, Beaudet, A.L., Sly, W.S. and Valle, D. New York, McGraw Hill. vol II: 4353-4392.

**Kuhlbrodt, K., B. Herbarth, E. Sock, I. Hermans-Borgmeyer and M. Wegner** (1998a). Sox10, a Novel Transcriptional Modulator in Glial Cells. *J. Neurosci.* **18**: 237-250.

**Kuhlbrodt, K., C. Schmidt, E. Sock, V. Pingault, N. Bondurand, M. Goossens and M. Wegner** (1998b). Functional analysis of *Sox10* mutations found in human Waardenburg- Hirschsprung patients. *J. Biol. Chem.* **273**: 23033-23038.

**LaBonne, C. and M. Bronner-Fraser** (1998). Neural crest induction in *Xenopus*: evidence for a two-signal model. *Development* **125**: 2403-2414.

**Lahav, R., C. Ziller, E. Dupin and N. M. Le Douarin** (1996). Endothelin 3 promotes neural crest cell proliferation and mediates a vast increase in melanocyte number in culture. *Proc. Natl. Acad. Sci. USA* **93**: 3892-3897.

**Lamers, C. H., J. W. Rombout and L. P. Timmermans** (1981). An experimental study on neural crest migration in *Barbus conchonius* (Cyprinidae, Teleostei), with special reference to the origin of the enteroendocrine cells. *J. Embryol. Exp. Morphol.* **62**: 309-323.

**Lane, P. W. and H. M. Liu** (1984). Association of megacolon with a new dominant spotting gene (*Dom*) in the mouse. *J. Heredity* **75**: 435-439.

**Lang, D., F. Chen, R. Milewski, J. Li, M. M. Lu and J. A. Epstein** (2000). Pax3 is required for enteric ganglia formation and functions with Sox10 to modulate expression of *c-ret*. *J. Clin. Invest.* **106**: 963-971.

**Laudet, V., D. Stehelin and H. Clevers** (1993). Ancestry and diversity of the HMG box superfamily. *Nucleic Acids Res.* **21**: 2493-2501.

**Le Douarin, N. M. and M. A. Teillet** (1974). Experimental analysis of the migration and differentiation of neuroblasts of the autonomic nervous system and of neurectodermal mesenchymal derivatives, using a biological cell marking technique. *Dev. Biol.* **41**: 162-184.

**Le Douarin, N. M. and C. Kalcheim** (1999). *The Neural Crest*. Cambridge, Press Syndicate of the University of Cambridge.

**Lefebvre, V., P. Li and B. De Crombrugghe** (1998). A new long form of Sox5 (L-Sox5), Sox6 and Sox9 are coexpressed in chondrogenesis and cooperatively activate the type II collagen gene. *Embo J.* **17**: 5718-5733.

**Lemke, G., E. Lamar and J. Patterson** (1988). Isolation and analysis of the gene encoding peripheral myelin protein zero. *Neuron* **1**: 73-83.

**Lister, J. A., J. Close and D. W. Raible** (2001). Duplicate *mitf* genes in zebrafish: complementary expression and conservation of melanogenic potential. *Dev. Biol.* **237**: 333-344.

**Lister, J. A., C. P. Robertson, T. Lepage, S. L. Johnson and D. W. Raible** (1999). *nacre* encodes a zebrafish microphthalmia-related protein that regulates neural-crest-derived pigment cell fate. *Development* **126**: 3757-3767.

**Love, J. J., X. Li, D. A. Case, K. Giese, R. Grosschedl and P. E. Wright** (1995). Structural basis for DNA bending by the architectural transcription factor LEF-1. *Nature* **376**: 791-795.

**Luo, R., M. An, B. L. Arduini and P. D. Henion** (2001). Specific pan-neural crest expression of zebrafish *crestin* throughout embryonic development. *Dev. Dyn.* **220**: 169-174.

**Mahakrishnan, A. and M. S. Srinivasan** (1980). Piebaldness with Hirschsprung's disease. *Arch Dermatol* **116**: 1102.

**Malicki, J., A. F. Schier, L. Solnica-Krezel, D. L. Stemple, S. C. Neuhauss, D. Y. Stainier, S. Abdelilah, Z. Rangini, F. Zwartkruis and W. Driever** (1996). Mutations affecting development of the zebrafish ear. *Development* **123**: 275-283.

**Marchant, L., C. Linker, P. Ruiz, N. Guerrero and R. Mayor** (1998). The inductive properties of mesoderm suggest that the neural crest cells are specified by a BMP gradient. *Dev. Biol.* **198**: 319-329.

**Marshall, O. J. and V. R. Harley** (2000). Molecular mechanisms of SOX9 action. *Mol. Genet. Metab.* **71**: 455-462.

**Masai, I., C. P. Heisenberg, K. A. Barth, R. Macdonald, S. Adamek and S. W. Wilson** (1997). *floating head* and *masterblind* regulate neuronal patterning in the roof of the forebrain. *Neuron* **18**: 43-57.

**Mayor, R., R. Morgan and M. G. Sargent** (1995). Induction of the prospective neural crest of *Xenopus*. *Development* **121**: 767-777.

**McGrew, L. L., S. Hoppler and R. T. Moon** (1997). Wnt and FGF pathways cooperatively pattern anteroposterior neural ectoderm in *Xenopus*. *Mech. Dev.* **69**: 105-114.

**Mertin, S., S. G. McDowall and V. R. Harley** (1999). The DNA-binding specificity of SOX9 and other SOX proteins. *Nucleic Acids Res.* **27**: 1359-1364.

**Morrison-Graham, K. and Y. Takahashi** (1993). Steel factor and c-kit receptor: from mutants to a growth factor system. *Bioessays* **15**: 77-83.

**Moury, J. D. and A. G. Jacobson** (1990). The origins of neural crest cells in the axolotl. *Dev. Biol.* **141**: 243-253.

**Mullins, M. C., M. Hammerschmidt, P. Haffter and C. Nuesslein-Volhard** (1994). Large-scale mutagenesis in the zebrafish: In search of genes controlling development in a vertebrate. *Curr. Biol.* **4**: 189-202.

**Mullins, M. C. and C. Nuesslein-Volhard** (1993). Mutational approaches to studying embryonic pattern formation in the zebrafish. *Curr. Opin. Genet. Dev.* **3**: 648-654.

**Murphy, M., K. Reid, D. E. Williams, S. D. Lyman and P. F. Bartlett** (1992). Steel factor is required for maintenance, but not differentiation, of melanocyte precursors in the neural crest. *Dev. Biol.* **153**: 396-401.

**Ng, L., S. Wheatley, G. E. Muscat, J. Conway-Campbell, J. Bowles, E. Wright, D. Bell, P. P. Tam, K. S. Cheah and P. Koopman** (1997). Sox9 binds DNA, activates transcription, and coexpresses with type II collagen during chondrogenesis in the mouse. *Dev. Biol.* **183**: 108-121.

**Nguyen, V. H., B. Schmid, J. Trout, S. A. Connors, M. Ekker and M. C. Mullins** (1998). Ventral and lateral regions of the zebrafish gastrula, including the neural crest progenitors, are established by a *bmp2b/swirl* pathway of genes. *Dev. Biol.* **199**: 93-110.

**Noden, D. M.** (1987). Interactions between cephalic neural crest and mesodermal populations. *Developmental and Evolutionary Aspects of the Neural Crest*. P. Maderson. New York, John Wiley and Sons: 89-119.

**Nuesslein-Volhard, C., E. Wieschaus and H. Kluding** (1984). Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster* I. zygotic loci on the second chromosome. *Roux's Arch. Dev. Biol.* **193**: 267-282.

**Odenthal, J. and C. Nuesslein-Volhard** (1998). *fork head* domain genes in zebrafish. *Dev. Genes Evol.* **208**: 245-258.

**Oetting, W. S.** (1998). Anatomy of pigment cell genes acting at the subcellular level. *The Pigmentary System. Physiology and Pathophysiology*. New York, Oxford University Press, Inc.: 231-249.

**Oliver, G., F. Loosli, R. Koster, J. Wittbrodt and P. Gruss** (1996). Ectopic lens induction in fish in response to the murine homeobox gene *Six3*. *Mech. Dev.* **60**: 233-239.

**Opdecamp, K., A. Nakayama, M. T. T. Nguyen, C. A. Hodgkinson, W. J. Pavan and H. Arnheiter** (1997). Melanocyte development *in vivo* and in neural crest cell cultures: Crucial dependence on the Mitf basic-helix-loop-helix-zipper transcription. *Development* **124**: 2377-2386.

**Paratore, C., D. E. Goerich, U. Suter, M. Wegner and L. Sommer** (2001). Survival and glial fate acquisition of neural crest cells are regulated by an interplay between the transcription factor Sox10 and extrinsic combinatorial signaling. *Development* **128**: 3949-3961.

**Parichy, D. M., J. F. Rawls, S. J. Pratt, T. T. Whitfield and S. L. Johnson** (1999). Zebrafish *sparse* corresponds to an orthologue of *c-kit* and is required for the

morphogenesis of a subpopulation of melanocytes, but is not essential for hematopoiesis or primordial germ cell development. *Development* **126**: 3425-3436.

**Pattyn, A., X. Morin, H. Cremer, C. Goridis and J. F. Brunet** (1999). The homeobox gene *Phox2b* is essential for the development of autonomic neural crest derivatives. *Nature* **399**: 366-370.

**Peirano, R. I., D. E. Goerich, D. Riethmacher and M. Wegner** (2000a). *Protein zero* gene expression is regulated by the glial transcription factor Sox10. *Mol. Cell. Biol.* **20**: 3198-3209.

**Peirano, R. I. and M. Wegner** (2000b). The glial transcription factor Sox10 binds to DNA both as monomer and dimer with different functional consequences. *Nucleic Acids Res.* **28**: 3047-3055.

**Peters, R., C. Y. King, E. Ukiyama, S. Falsafi, P. K. Donahoe and M. A. Weiss** (1995). An *SRY* mutation causing human sex reversal resolves a general mechanism of structure-specific DNA recognition: application to the four-way DNA junction. *Biochemistry* **34**: 4569-4576.

**Pevny, L. H. and R. Lovell-Badge** (1997). *Sox* genes find their feet. *Curr. Opin. Gen. Dev.* 7: 338-344.

**Piccolo, S., Y. Sasai, B. Lu and E. M. De Robertis** (1996). Dorsoventral patterning in *Xenopus*: inhibition of ventral signals by direct binding of chordin to BMP-4. *Cell* **86**: 589-598.

**Pingault, V., N. Bondurand, K. Kuhlbrodt, D. E. Goerich, M. O. Prehu, A. Puliti, B. Herbarth, I. Hermans-Borgmeyer, E. Legius, G. Matthijs, J. Amiel, S. Lyonnet, I. Ceccherini, G. Romeo, J. C. Smith, A. P. Read, M. Wegner and M. Goossens** (1998a). *SOX10* mutations in patients with Waardenburg-Hirschsprung disease. *Nat. Genet.* **18**: 171-173.

**Pingault, V., B. Herbarth, N. Bondurand, K. Kuhlbrodt, I. Hermans-Borgmeyer, A. Puliti, N. Lemort, M. O. Prehu, M. Goossens and M. Wegner** (1998b). Mutation of the *Sry*-related *Sox10* gene in *dominant megacolon*, a mouse model for human Hirschsprung disease. *Europ. J. Hum. Genet.* **6**: 4153.

**Pingault, V., A. Puliti, M. Prehu, A. Samadi, N. Bondurand and M. Goossens** (1997). Human Homology and Candidate Genes for the *Dominant megacolon* Locus, a Mouse Model for Hirschsprung Disease. *Genomics* **39**: 86-89.

**Postlethwait, J., Amores, A., Force, A. and Yan, Yi-Lin** (1999). The Zebrafish Genome. *Meth. Cell Biol.*, Academic Press. **60:** 149-163.

**Postlethwait, J. H., S. L. Johnson, C. N. Midson, W. S. Talbot, M. Gates, E. W. Ballinger, D. Africa, R. Andrews, T. Carl, J. S. Eisen, S. Horne, C. B. Kimmel, M. Hutchinson, M. Johnson and A. Rodriguez** (1994). A genetic-linkage map for the zebrafish. *Science* **264**: 699-703.

**Postlethwait, J. H., Y. L. Yan, M. A. Gates, S. Horne, A. Amores, A. Brownlie, A. Donovan, E. S. Egan, A. Force, Z. Y. Gong, C. Goutel, A. Fritz, R. Kelsh, E. Knapik, E. Liao, B. Paw, D. Ransom, A. Singer, M. Thomson, T. S. Abduljabbar, P. Yelick, J. Beier, J. S. Joly, D. Larhammar, F. Rosa, M. Westerfield, L. I. Zon, S. L. Johnson, W. S. Talbot and M. Ekker** (1998). Vertebrate genome evolution and the zebrafish gene map (vol 18, pg 345, 1998). *Nat. Genet.* **19**: 303.

**Potterf, S. B., M. Furumura, K. J. Dunn, H. Arnheiter and W. J. Pavan** (2000). Transcription factor hierarchy in Waardenburg syndrome: regulation of *MITF* expression by SOX10 and PAX3. *Hum. Genet.* **107**: 1-6.

**Potterf, S. B., R. Mollaaghababa, L. Hou, E. M. Southard-Smith, T. J. Hornyak, H. Arnheiter and W. J. Pavan** (2001). Analysis of SOX10 function in neural crest-derived melanocyte development: SOX10-dependent transcriptional control of dopachrome tautomerase. *Dev. Biol.* **237**: 245-257.

**Puffenberger, E. G., K. Hosoda, S. S. Washington, K. Nakao, D. deWit, M. Yanagisawa and A. Chakravarti** (1994a). A missense mutation of the *endothelin-B receptor* gene in multigenic Hirschsprung's disease. *Cell* **79**: 1257-1266.

**Puffenberger, E. G., E. R. Kauffman, S. Bolk, T. C. Matise, S. S. Washington, M. Angrist, J. Weissenbach, K. L. Garver, M. Mascari, R. Ladda** (1994b). Identity-by-descent and association mapping of a recessive gene for Hirschsprung disease on human chromosome 13q22. *Hum. Mol. Genet.* **3**: 1217-1225.

**Puliti, A., V. Poirier, M. Goossens and M. Simonneau** (1996). Neuronal defects in genotyped dominant megacolon (*Dom*) mouse embryos, a model for Hirschsprung disease. *Neuroreport* **7**: 489-492.

**Pusch, C., E. Hustert, D. Pfeifer, P. Suedbeck, R. Kist, B. Roe, Z. L. Wang, R. Balling, N. Blin and G. Scherer** (1998). The *SOX10/Sox10* gene from human and mouse: sequence, expression, and transactivation by the encoded HMG domain transcription factor. *Hum. Genet.* **103**: 115-123.

**Raible, D. W. and J. S. Eisen** (1994). Restriction of neural crest cell fate in the trunk of the embryonic zebrafish. *Development* **120**: 495-503.

**Raible, D. W. and J. S. Eisen** (1995). Lateral specification of cell fate during vertebrate development. *Curr. Opin. Gen. Dev.* **5**: 444-449.

**Raible, D. W. and J. S. Eisen** (1996). Regulative interactions in zebrafish neural crest. *Development* **122**: 501-507.

**Raible, D. W., A. Wood, W. Hodsdon, P. D. Henion, J. A. Weston and J. S. Eisen** (1992). Segregation and early dispersal of neural crest cells in the embryonic zebrafish. *Dev. Dyn.* **195**: 29-42.

**Raven, C. P. and J. Kloos** (1945). Induction by medial and lateral pieces of the archenteron roof with special reference to the determination of the neural crest. *Acta neerl. Morphol.* **5**: 348-362.

**Rickmann, M., J. W. Fawcett and R. J. Keynes** (1985). The migration of neural crest cells and the growth of motor neurons through the rostral half of the chick somite. *J. Embryol. Exp. Morphol.* **90**: 437-455.

**Riehl, R. and H. A. Baensch** (1996). *Aquarium Atlas*. Melle, Germany, Baensch, H.A.

**Rimini, R., M. Beltrame, F. Argenton, D. Szymczak, F. Cotelli and M. E. Bianchi** (1999). Expression patterns of zebrafish *sox11A*, *sox11B* and *sox21*. *Mech. Dev.* **89**: 167-171.

**Rollhauser-ter Horst, J.** (1980). Neural crest replaced by gastrula ectoderm in amphibia. Effect on neurulation, CNS, gills and limbs. *Anat. Embryol.* **160**: 203-211.

**Romeo, G., P. Ronchetto, Y. Luo, V. Barone, M. Seri, I. Ceccherini, B. Pasini, R. Bocciardi, M. Lerone, H. Kaariainen** (1994). Point mutations affecting the tyrosine kinase domain of the *RET* proto-oncogene in Hirschsprung's disease. *Nature* **367**: 377-378.

**Roose, J., W. Korver, E. Oving, A. Wilson, G. Wagenaar, M. Markman, W. Lamers and H. Clevers** (1998). High expression of the HMG box factor *sox-13* in arterial walls during embryonic development. *Nucleic Acids Res.* **26**: 469-476.

**Russell, W. L.** (1947). *Splotch*, a new mutation in the house mouse, *Mus musculus*. *Genetics* **32**: 102.

**Sambrook, J., E. F. Fritsch and T. Maniatis** (1989). *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbour, Cold Spring Harbour Laboratory Press.

**Sanders, P.** (2000). Quantitative analysis of neural crest defects in *colourless* mutant embryos. Bath, Department of Biology and Biochemistry, University of Bath.

**Schepers, G. E., M. Bullejos, B. M. Hosking and P. Koopman** (2000). Cloning and characterisation of the *Sry*-related transcription factor gene *Sox8*. *Nucleic Acids Res.* **28**: 1473-1480.

**Schilling, T. F. and C. B. Kimmel** (1994). Segment and cell-type lineage restrictions during pharyngeal arch development in the zebrafish. *Development* **120**: 483-494.

**Schneider, C., H. Wicht, J. Enderich, M. Wegner and H. Rohrer** (1999). Bone morphogenetic proteins are required *in vivo* for the generation of sympathetic neurons. *Neuron* **24**: 861-870.

**Schuchardt, A., V. D'Agati, L. Larsson-Blomberg, F. Costantini and V. Pachnis** (1994). Defects in the kidney and enteric nervous system of mice lacking the tyrosine kinase receptor Ret. *Nature* **367**: 380-383.

**Selleck, M. A. J. and M. Bronner-Fraser** (1995). Origins of the avian neural crest - The role of neural plate-epidermal interactions. *Development* **121**: 525-538.

**Serbedzija, G. N., S. E. Fraser and M. Bronner-Fraser** (1990). Pathways of trunk neural crest cell migration in the mouse embryo as revealed by vital dye labelling. *Development* **108**: 605-612.

**Shah, K. N., S. J. Dalal, M. P. Desai, P. N. Sheth, N. C. Joshi and L. M. Ambani** (1981). White forelock, pigmentary disorder of irides, and long segment Hirschsprung disease: possible variant of Waardenburg syndrome. *J. Pediatr.* **99**: 432-435.

**Shin, M. K., J. M. Levorse, R. S. Ingram and S. M. Tilghman** (1999). The temporal requirement for endothelin receptor-B signalling during neural crest development. *Nature* **402**: 496-501.

**Sieber-Blum, M.** (1990). Mechanism of neural crest diversification. *Comments Dev. Neurobiol.* **1**: 225-249.

**Smith, M., A. Hickman, D. Amanze, A. Lumsden and P. Thorogood** (1994). Trunk neural crest origin of caudal fin mesenchyme in the zebrafish *Brachydanio rerio*. *Proc. R. Soc. Lond. B* **256**: 137-145.

**Smith, W. C. and R. M. Harland** (1992). Expression cloning of noggin, a new dorsalizing factor localized to the Spemann organizer in *Xenopus* embryos. *Cell* **70**: 829-840.

**Sommer, L., Q. Ma and D. J. Anderson** (1996). *neurogenins*, a Novel Family of *atonal*-Related bHLH Transcription Factors, Are Putative Mammalian Neuronal Determination Genes That Reveal Progenitor Cell Heterogeneity in the Developing CNS and PNS. *Mol. Cell. Neurosci.* **8**: 221-241.

**Southard-Smith, E. M., M. Angrist, J. S. Ellison, R. Agarwala, A. D. Baxevanis, A. Chakravarti and W. J. Pavan** (1999). The *Sox10*$^{Dom}$ mouse: Modeling the genetic variation of Waardenburg-Shah (WS4) syndrome. *Genome Research* **9**: 215-225.

**Southard-Smith, E. M., L. Kos and W. J. Pavan** (1998). *Sox10* mutation disrupts neural crest development in DOM Hirschsprung mouse model. *Nat. Genet.* **18**: 60-64.

**Spritz, R. A.** (1998a). Piebaldism, Waardenburg Syndrome, and Related Genetic Disorders-Molecular and Genetic Aspects. *The Pigmentary System. Physiology and Pathophysiology*. New York, Oxford University Press, Inc.: 207-215.

**Spritz, R. A. and J.-P. Ortonne** (1998b). Genetic Hypomelanoses: Disorders characterised by congenital depigmentation. *The Pigmentary System. Physiology and Pathophysiology*. New York, Oxford University Press, Inc.: 505-510.

**Steel, K. and C. Barkway** (1989). Another role for melanocytes; their importance for normal stria vascularis development in the mammalian inner ear. *Development* **107**: 453-463.

**Steel, K. P., D. R. Davidson and I. J. Jackson** (1992). *TRP-2/DCT*, a new early melanoblast marker, shows that steel growth factor (c-kit ligand) is a survival factor. *Development* **115**: 1111-1119.

**Steingrimsson, E., K. J. Moore, M. L. Lamoreux, A. R. Ferre-D'Amare, S. K. Burley, D. C. Zimring, L. C. Skow, C. A. Hodgkinson, H. Arnheiter and N. G. Copeland** (1994). Molecular basis of mouse *microphthalmia* (*mi*) mutations helps explain their developmental and phenotypic consequences. *Nat. Genet.* **8**: 256-263.

**Stolt, C. C., S. Rehberg, M. Ader, P. Lommes, D. Riethmacher, M. Schachner, U. Bartsch and M. Wegner** (2002). Terminal differentiation of myelin-forming oligodendrocytes depends on the transcription factor Sox10. *Gen. Dev.* **16**: 165-170.

**Strimmer, K. and A. vonHaeseler** (1996). Quartet puzzling: A quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**: 964-969.

**Suedbeck, P. and G. Scherer** (1997). Two independent nuclear localization signals are present in the DNA-binding high-mobility group domains of SRY and SOX9. *J. Biol. Chem.* **272**: 27848-27852.

**Suedbeck, P., M. L. Schmitz, P. A. Baeuerle and G. Scherer** (1996). Sex reversal by loss of the C-terminal transactivation domain of human SOX9. *Nat. Genet.* **13**: 230-232.

**Tachibana, M., K. Takeda, Y. Nobukuni, K. Urabe, J. E. Long, K. A. Meyers, S. A. Aaronson and T. Miki** (1996). Ectopic expression of *MITF*, a gene for Waardenburg syndrome type 2, converts fibroblasts to cells with melanocyte characteristics. *Nat. Genet.* **14**: 50-54.

**Tani, M., N. ShindoOkada, Y. Hashimoto, T. Shiroishi, S. Takenoshita, Y. Nagamachi and J. Yokota** (1997). Isolation of a novel *Sry*-related gene that is expressed in high-metastatic K-1735 murine melanoma cells. *Genomics* **39**: 30-37.

**Taraviras, S., C. V. MarcosGutierrez, P. Durbec, H. Jani, M. Grigoriou, M. Sukumaran, L. C. Wang, M. Hynes, G. Raisman and V. Pachnis** (1999). Signalling by the RET receptor tyrosine kinase and its role in the development of the mammalian enteric nervous system. *Development* **126**: 2785-2797.

**Tassabehji, M., V. E. Newton and A. P. Read** (1994). Waardenburg syndrome type 2 caused by mutations in the human microphthalmia (*MITF*) gene. *Nat. Genet.* **8**: 251-255.

**Tassabehji, M., A. P. Read, V. E. Newton, R. Harris, R. Balling, P. Gruss and T. Strachan** (1992). Waardenburg's syndrome patients have mutations in the human homologue of the *Pax-3* paired box gene. *Nature* **355**: 635-636.

**Thisse, C., B. Thisse and J. H. Postlethwait** (1995). Expression of *snail2*, a 2nd member of the zebrafish *snail* family, in cephalic mesendoderm and presumptive neural crest of wild-type and *spadetail* mutant embryos. *Dev. Biol.* **172**: 86-99.

**Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins** (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**: 4876-4882.

**Uchikawa, M., Y. Kamachi and H. Kondoh** (1999). Two distinct subgroups of Group B *Sox* genes for transcriptional activators and repressors: their expression during embryonic organogenesis of the chicken. *Mech. Dev.* **84**: 103-120.

**Uwanogho, D., M. Rex, E. J. Cartwright, G. Pearl, C. Healy, P. J. Scotting and P. T. Sharpe** (1995). Embryonic expression of the chicken *Sox2*, *Sox3* and *Sox11* genes suggests an interactive role in neuronal development. *Mech. Dev.* **49**: 23-36.

**van Houte, L. P., V. P. Chuprina, M. van der Wetering, R. Boelens, R. Kaptein and H. Clevers** (1995). Solution structure of the sequence-specific HMG box of the lymphocyte transcriptional activator Sox-4. *J Biol Chem* **270**: 30516-30524.

**Vriz, S. and R. Lovell-Badge** (1995). The zebrafish Zf-Sox19 protein: a novel member of the Sox family which reveals highly conserved motifs outside the DNA-binding domain. *Gene* **153**: 275-276.

**Waardenburg, P. J.** (1951). A new syndrome combining developmental anomalies of the eyelids, eyebrows and nose root with pigmentary defects of the iris and head hair and with congenital deafness. *Am. J. Hum. Genet.* **3**: 195-253.

**Wagner, T., J. Wirth, J. Meyer, B. Zabel, M. Held, J. Zimmer, J. Pasantes, F. D. Bricarelli, J. Keutel and E. Hustert** (1994). Autosomal sex reversal and campomelic dysplasia are caused by mutations in and around the *SRY*-related gene *SOX9*. *Cell* **79**: 1111-1120.

**Watanabe, A., K. Takeda, B. Ploplis and M. Tachibana** (1998). Epistatic relationship between Waardenburg syndrome genes *MITF* and *PAX3*. *Nat. Genet.* **18**: 283-286.

**Watanabe, K., K. Takeda, Y. Katori, K. Ikeda, T. Oshima, K. Yasumoto, H. Saito, T. Takasaka and S. Shibahara** (2000). Expression of the *Sox10* gene during mouse inner ear development. *Brain Res. Mol. Brain Res.* **84**: 141-145.

**Weinstein, D. C. and A. Hemmati-Brivanlou** (1997). Neural induction in *Xenopus* laevis: evidence for the default model. *Curr. Opin. Neurobiol.* **7**: 7-12.

**Werner, M. H. and S. K. Burley** (1997). Architectural transcription factors: proteins that remodel DNA. *Cell* **88**: 733-736.

**Werner, M. H., J. R. Huth, A. M. Gronenborn and G. M. Clore** (1995). Molecular basis of human 46X,Y sex reversal revealed from the three- dimensional solution structure of the human SRY-DNA complex. *Cell* **81**: 705-714.

**Westerfield, M.** (1995). *The Zebrafish Book. A guide for the laboratory use of zebrafish (Danio rerio)*. Eugene, Oregon, University of Oregon Press.

**Whitfield, T. T., M. Granato, F. J. M. vanEeden, U. Schach, M. Brand, M. Furutani-Seiki, P. Haffter, M. Hammerschmidt, C. P. Heisenberg, Y. J. Jiang, D. A. Kane, R. N. Kelsh, M. C. Mullins, J. Odenthal and C. Nuesslein-Volhard** (1996). Mutations affecting development of the zebrafish inner ear and lateral line. *Development* **123**: 241-254.

**Wieschaus, E., C. Nuesslein-Volhard and G. Juergens** (1984). Mutations affecting the pattern of the larval cuticle in *Drosophila melanogaster* III. Zygotic loci on

the X-chromosome and fourth chromosome. *Roux's Arch. Dev. Biol.* **193**: 296-307.

**Wilson, P. A. and A. Hemmati-Brivanlou** (1995). Induction of epidermis and inhibition of neural fate by Bmp-4. *Nature* **376**: 331-333.

**Wolffe, A. P.** (1994). Architectural transcription factors. *Science* **264**: 1100-1101.

**Wright, E., M. R. Hargrave, J. Christiansen, L. Cooper, J. Kun, T. Evans, U. Gangadharan, A. Greenfield and P. Koopman** (1995). The *Sry*-related gene *Sox9* is expressed during chondrogenesis in mouse embryos. *Nat. Genet.* **9**: 15-20.

**Wright, E. M., B. Snopek and P. Koopman** (1993). Seven new members of the *Sox* gene family expressed during mouse development. *Nucleic Acids Res* **21**: 744.

**Yang, G. C., D. Croaker, A. L. Zhang, P. Manglick, T. Cartmill and D. Cass** (1998). A dinucleotide mutation in the *endothelin-B receptor* gene is associated with lethal white foal syndrome (LWFS); a horse variant of Hirschsprung disease. *Hum. Mol. Genet.* **7**: 1047-1052.

**Yuan, H., N. Corbi, C. Basilico and L. Dailey** (1995). Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Gen. Dev.* **9**: 2635-2645.

**Zimmerman, L. B., J. M. De Jesus-Escobar and R. M. Harland** (1996). The Spemann organizer signal noggin binds and inactivates bone morphogenetic protein 4. *Cell* **86**: 599-606.

**Zlotogora, J., I. Lerer, S. Bar-David, Z. Ergaz and D. Abeliovich** (1995). Homozygosity for Waardenburg syndrome. *Am. J. Hum. Genet.* **56**: 1173-1178.

| 24hpf-WT | LP | MP | pre | LP | MP | pre | LP | MP | pre |
|---|---|---|---|---|---|---|---|---|---|
| som 1-5 | 3 | 22 | 27 | 3 | 13 | 15 | 5 | 13 | 36 |
| som 6-10 | 1 | 34 | 48 | / | 17 | 22 | 2 | 21 | 50 |
| som 11-15 | / | 21 | 53 | / | 22 | 44 | / | 25 | 75 |
| som 16-20 | / | 8 | 66 | / | 15 | 41 | / | 22 | 60 |

| 24hpf-WT | LP | MP | pre | LP | MP | pre |
|---|---|---|---|---|---|---|
| som 1-5 | 6 | 13 | 38 | 6 | 17 | 45 |
| som 6-10 | 2 | 24 | 40 | 4 | 19 | 59 |
| som 11-15 | / | 14 | 46 | 1 | 20 | 72 |
| som 16-20 | / | 16 | 43 | / | 16 | 60 |

| 24hpf-cls | LP | MP | pre | LP | MP | pre | LP | MP | pre |
|---|---|---|---|---|---|---|---|---|---|
| som 1-5 | 2 | 16 | 29 | 2 | 17 | 32 | / | 22 | 52 |
| som 6-10 | / | 22 | 37 | / | 21 | 43 | / | 23 | 71 |
| som 11-15 | / | 15 | 45 | / | 16 | 61 | / | 11 | 64 |
| som 16-20 | / | 13 | 65 | / | 14 | 45 | / | 8 | 56 |

| 24hpf-cls | LP | MP | pre | LP | MP | pre |
|---|---|---|---|---|---|---|
| som 1-5 | 2 | 17 | 48 | 1 | 17 | 51 |
| som 6-10 | / | 22 | 81 | / | 28 | 81 |
| som 11-15 | / | 19 | 84 | / | 19 | 79 |
| som 16-20 | / | 6 | 74 | / | 14 | 69 |

**Appendix 4.1:** Counts of *sox10*-positive premigratory and migratory cells. *sox10*-positive cells (sox10+) and differentiated melanophores (M+) were counted in the premigratory area (pre), on the lateral pathway (LP) and the medial pathway (MP) within muscle segments (som) 1-5, 6-10, 11-15 and 16-20 of wild-type (WT) and *cls* mutant embryos (cls). Counts were performed on 5 WT and *cls* embryos of each of 24hpf, 30hpf and 35hpf stages.

| | LP | | MP | | | LP | | MP | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 30hpf-WT | M+ | sox10+ | M+ | sox10+ | pre | M+ | sox10+ | M+ | sox10+ | pre |
| som 1-5 | 3 | / | 2 | 25 | 4 | 4 | / | 2 | 25 | 4 |
| som 6-10 | 2 | / | 6 | 26 | 7 | 1 | / | 9 | 18 | 6 |
| som 11-15 | 3 | / | 7 | 24 | 12 | 2 | / | 10 | 26 | 7 |
| som 16-20 | / | / | 4 | 23 | 8 | 3 | / | 10 | 14 | 10 |

| | LP | | MP | | | LP | | MP | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 30hpf-WT | M+ | sox10+ | M+ | sox10+ | pre | M+ | sox10+ | M+ | sox10+ | pre |
| som 1-5 | 3 | / | 6 | 29 | 2 | 2 | / | 3 | 25 | 3 |
| som 6-10 | 6 | / | 10 | 35 | 7 | 3 | / | 7 | 20 | 5 |
| som 11-15 | 2 | / | 15 | 32 | 12 | 3 | / | 9 | 26 | 11 |
| som 16-20 | 1 | / | 5 | 27 | 11 | 1 | / | 5 | 14 | 7 |

| | LP | | MP | | |
|---|---|---|---|---|---|
| 30hpf-WT | M+ | sox10+ | M+ | sox10+ | pre |
| som 1-5 | 1 | / | 2 | 27 | 5 |
| som 6-10 | 1 | / | 12 | 18 | 8 |
| som 11-15 | 2 | / | 12 | 27 | 5 |
| som 16-20 | 1 | / | 10 | 22 | 6 |

| | LP | | MP | | | LP | | MP | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 30hpf-cls | M+ | sox10+ | M+ | sox10+ | pre | M+ | sox10+ | M+ | sox10+ | pre |
| som 1-5 | / | 1 | / | 25 | 10 | / | 1 | / | 30 | 18 |
| som 6-10 | / | 1 | / | 26 | 24 | / | / | / | 33 | 27 |
| som 11-15 | / | / | / | 22 | 32 | / | / | / | 30 | 50 |
| som 16-20 | / | / | / | 20 | 33 | / | / | / | 16 | 62 |

| | LP | | MP | | | LP | | MP | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 30hpf-cls | M+ | sox10+ | M+ | sox10+ | pre | M+ | sox10+ | M+ | sox10+ | pre |
| som 1-5 | / | / | / | 31 | 14 | / | / | / | 31 | 22 |
| som 6-10 | / | / | / | 30 | 18 | / | / | / | 35 | 23 |
| som 11-15 | / | / | / | 24 | 63 | / | / | / | 29 | 56 |
| som 16-20 | / | / | / | 15 | 65 | / | / | / | 24 | 64 |

| | LP | | MP | | |
|---|---|---|---|---|---|
| 30hpf-cls | M+ | sox10+ | M+ | sox10+ | pre |
| som 1-5 | / | / | / | 28 | 9 |
| som 6-10 | / | / | / | 29 | 18 |
| som 11-15 | / | / | / | 21 | 39 |
| som 16-20 | / | / | / | 11 | 42 |

| | LP | | MP | | | LP | | MP | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 35hpf-WT | M+ | sox10+ | M+ | sox10+ | pre | M+ | sox10+ | M+ | sox10+ | pre |
| som 1-5 | 3 | / | 3 | 16 | 7 | 5 | / | 2 | 17 | 5 |
| som 6-10 | 1 | / | 3 | 11 | 7 | 3 | / | 7 | 17 | 8 |
| som 11-15 | / | / | 7 | 7 | 13 | 6 | / | 5 | 20 | 10 |
| som 16-20 | 1 | / | 5 | 13 | 9 | / | / | 8 | 20 | 13 |

| | LP | | MP | | | LP | | MP | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 35hpf-WT | M+ | sox10+ | M+ | sox10+ | pre | M+ | sox10+ | M+ | sox10+ | pre |
| som 1-5 | 1 | / | 2 | 19 | 8 | 6 | / | 1 | 26 | 8 |
| som 6-10 | 2 | / | 6 | 18 | 9 | 3 | / | 7 | 20 | 11 |
| som 11-15 | 2 | / | 6 | 15 | 12 | 3 | / | 2 | 23 | 8 |
| som 16-20 | 2 | / | 8 | 23 | 11 | 3 | / | 7 | 20 | 8 |

| | LP | | MP | | |
|---|---|---|---|---|---|
| 35hpf-WT | M+ | sox10+ | M+ | sox10+ | pre |
| som 1-5 | 3 | / | 3 | 17 | 8 |
| som 6-10 | 4 | / | 4 | 23 | 8 |
| som 11-15 | 4 | / | 8 | 20 | 10 |
| som 16-20 | 1 | / | 8 | 26 | 11 |

| | LP | | MP | | | LP | | MP | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 35hpf-cls | M+ | sox10+ | M+ | sox10+ | pre | M+ | sox10+ | M+ | sox10+ | pre |
| som 1-5 | / | / | / | 21 | 29 | / | / | / | 26 | 23 |
| som 6-10 | / | / | / | 26 | 32 | / | / | / | 21 | 26 |
| som 11-15 | / | 1 | / | 20 | 46 | / | / | / | 22 | 39 |
| som 16-20 | / | / | / | 19 | 36 | / | / | / | 20 | 30 |

| | LP | | MP | | | LP | | MP | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 35hpf-cls | M+ | sox10+ | M+ | sox10+ | pre | M+ | sox10+ | M+ | sox10+ | pre |
| som 1-5 | / | 1 | / | 14 | 29 | / | / | / | 20 | 21 |
| som 6-10 | / | / | / | 16 | 39 | / | 1 | / | 23 | 36 |
| som 11-15 | / | / | / | 12 | 46 | / | / | / | 14 | 37 |
| som 16-20 | / | / | / | 12 | 38 | / | / | / | 13 | 28 |

| | LP | | MP | | |
|---|---|---|---|---|---|
| 35hpf-cls | M+ | sox10+ | M+ | sox10+ | pre |
| som 1-5 | / | 2 | / | 16 | 20 |
| som 6-10 | / | 1 | / | 21 | 29 |
| som 11-15 | / | / | / | 14 | 35 |
| som 16-20 | / | / | / | 18 | 32 |

| Allele/batch | No. of embryos | Vol. resuspended (µl) | No. embryos/µl |
|---|---|---|---|
| *tw1* | 38 | 20 | 1.9 |
| *tw2*/1 | 123 | 40 | 3.1 |
| *tw2*/2 | 117 | 40 | 2.9 |
| *t3*/1 | 113 | 40 | 2.8 |
| *t3*/2 | 63 | 40 | 1.6 |
| *tw11* | 49 | 30 | 1.6 |
| *ty22f* | 15 | 20 | 0.8 |
| *m618*/1 | 83 | 40 | 2.1 |
| *m618*/2 | 46 | 30 | 1.5 |

**Appendix 6.1**: Preparation of total RNA from *cls* mutant alleles

The table summarises the number of homozygous *cls* mutant embryos that were collected from each allele and the volume DEPC-treated water used to resuspend the RNA pellet. For alleles *tw2*, *t3* and *m618* enough mutant embryos were available to split them into two batches. The last column (No. embryos/µl) attempts to give a rough estimation of relative differences in resulting RNA concentrations of each sample by listing the number of embryos RNA was isolated from per µl final volume of resulting RNA sample. No., number; Vol., volume.

**Appendix 7.1:** Comparison of rescue of homozygous *cls*$^{m618}$ mutant embryos by ectopic expression of wild-type and mutant Sox10 proteins under varied heatshock and microinjection regimes

(A) Various batches of embryos (WT-A to WT-G and WT-1 to WT-37) from a heterozygous *cls*$^{m618}$ cross were injected with either 12pg, 25pg, 50pg or 120pg of hs>*cls*$^{WT}$. (B, C) Similarly, 50pg or 120pg of hs>*cls*$^{m618}$ (B) or hs>*cls*$^{tw2}$ (C) were injected into various batches of embryos (M-1 to M-11 for hs>*cls*$^{m618}$ and T-1 to T-18 for hs>*cls*$^{tw2}$) from a heterozygous *cls*$^{m618}$ cross. Batches grouped by a bracket were injected on the same day. (A-C) All batches were heatshocked as indicated in column labelled "heatshock". The column "uninjected/ % survived" in all tables list the total number of uninjected embryos and the percentage of uninjected embryos that had survived after 48hours. In the columns "injected/ % survived", the total number of injected siblings and the percentage of injected embryos that had survived after 48hours are listed. The number of *cls* embryos within these alive and injected embryos is shown together with the percentage of rescued *cls* embryos within the surviving *cls* siblings in the columns "*cls*/ % rescued". The last columns list the mean number of melanophores (M) with wild-type morphology observed per rescued *cls* embryo. nd, not determined; "-", no rescued melanophores.

A

## hs>*cls*$^{WT}$ injections (50pg) with two heatshocks:

( ≥ 1 dark spidery melanophore counts as rescued)

| *hs>cls*$^{WT}$ | 1st/2nd heatshock | uninjected/ % survived | injected / % survived | *cls* / % rescued | mean number of M |
|---|---|---|---|---|---|
| WT-A | 11.7hpf/18hpf | nd | 260/31.5% | 20/50% | 6.6 |
| WT-B | 10hpf/16hpf | nd | 80/50% | 4/0% | - |
| WT-C | 18hpf/24hpf | nd | 35/17% | 0/0% | - |
| WT-D | 18hpf/24hpf | nd | 28/25% | 1/0% | - |
| WT-E | 18hpf/24hpf | nd | 33/30.3% | 2/100% | nd |
| WT-F | 13hpf/24hpf | 30/97% | 157/90% | 20/35% | 12.3 |
| WT-G | 11.3hpf/22hpf | 29/86% | 213/85% | 28/61% | 7.4 |
| WT-3 | 11hpf/23hpf | 14/64% | 91/56% | 12/42% | 12.2 |
| WT-4 | 10.5hpf/22.5hpf | nd | 19/79% | 3/100% | 17.6 |
| WT-5 | 10hpf/22hpf | 15/93% | 218/69% | 27/37% | 14.7 |

## hs>*cls*$^{WT}$ injections (50pg) without heatshock:

| *hs>cls*$^{WT}$ | heatshock | uninjected/ % survived | injected / % survived | *cls* / % rescued | mean number of M |
|---|---|---|---|---|---|
| WT-6 | / | 10/100% | 64/57.8% | 10/0% | - |
| WT-7 | / | 22/45% | 126/25.4% | 7/28.6% | 2.0 |
| WT-8 | / | 11/0% | 120/11.7% | 1/0% | - |
| WT-9 | / | 20/70% | 119/24.4% | 4/75% | 3.3 |
| WT-10 | / | 38/76.3% | 178/43.8% | 14/57% | 4.0 |
| WT-11 | / | 17/47% | 155/28.4% | 12/33.3% | 7.0 |

## hs>*cls*$^{WT}$ injections (25pg) without heatshock:

| hs>*cls*$^{WT}$ | heatshock | uninjected/ % survived | injected / % survived | *cls* / % rescued | mean number of M |
|---|---|---|---|---|---|
| WT-25 | / | 12/83.3% | 123/70.7% | 28/10.7% | 3.0 |
| WT-26 | / | 11/63.6% | 81/40.7% | 8/12.5% | 1.0 |
| WT-27 | / | 16/100% | 188/86.7% | 43/14% | 5.2 |
| WT-28 | / | 106/90.6% | 224/77.2% | 36/5.5% | 1.5 |
| WT-29 | / | nd | 33/84.8% | 7/14.3% | 2.0 |

## hs>*cls*$^{WT}$ injections (25pg) and only one heatshock:

| hs>*cls*$^{WT}$ | heatshock | uninjected/ % survived | injected / % survived | *cls* / % rescued | mean number of M |
|---|---|---|---|---|---|
| WT-12 | 15.5hpf | 4/50% | 13/23.1% | 0/0% | - |
| WT-13 | 16hpf | 9/100% | 104/75% | 21/23.8% | 6.2 |
| WT-14 | 18.5hpf | 18/100% | 238/91.2% | 51/31.4% | 6.3 |
| WT-15 | 15.5hpf | 20/100% | 111/87.4% | 19/21.1% | 1.3 |
| WT-16 | 15.5hpf | 13/100% | 44/88.6% | 8/12.5% | 2.0 |

## hs>*cls*$^{WT}$ injections (50pg) and only one heatshock:

| hs>*cls*$^{WT}$ | heatshock | uninjected/ % survived | injected / % survived | *cls* / % rescued | mean number of M |
|---|---|---|---|---|---|
| WT-17 | 15.5hpf | 15/100% | 175/70.9% | 9/77.8% | 31.9 |
| WT-18 | 15.5hpf | 16/100% | 246/77.7% | 24/54.2% | 11.6 |
| WT-19 | 18hpf | nd | 22/0% | 0/0% | - |
| WT-20 | 18.5hpf | 20/80% | 134/41% | 13/61.5% | 2.4 |
| WT-21 | 18.5hpf | 16/50% | 94/37.2% | 6/50% | 3.6 |
| WT-22 | 18hpf | 32/31.3% | 234/26.5% | 16/25% | 9.0 |
| WT-23 | 18hpf | 17/29.4% | 66/22.7% | 1/0% | - |
| WT-24 | 18.5hpf | 9/100% | 72/40.3% | 7/85.7% | 9.0 |

## hs>*cls*$^{WT}$ injections (12pg) with one early heatshock:

| hs>*cls*$^{WT}$ | heatshock | uninjected/ % survived | injected/ % survived | *cls* / % rescued | mean number of M |
|---|---|---|---|---|---|
| WT-30 | 15-16hpf | 15/66.7% | 199/14.1% | 5/0% | - |
| WT-31 | 15-16hpf | 13/0% | 69/39.1% | 9/11.1% | 1.0 |
| WT-32 | 15-16hpf | 10/50% | 162/22.8% | 14/14.3% | 1.0 |
| WT-33 | 15-16hpf | / | 76/5.3% | 1/0% | - |
| WT-34 | 15-16hpf | 10/20% | 93/4.3% | 0/0% | - |
| WT-38 | 15-16hpf | nd | 133/71.4% | 20/5% | 2.0 |
| WT-39 | 15-16hpf | 15/93.3% | 276/44.6% | 33/12.1% | 11.8 |

## hs>*cls*$^{WT}$ injections (120pg) with one early heatshock:

| hs>*cls*$^{WT}$ | heatshock | uninjected/ % survived | injected / % survived | *cls* / % rescued | mean number of M |
|---|---|---|---|---|---|
| WT-35 | 15-16hpf | 15/53.3% | 191/39.3% | 13/76.9% | 14.1 |
| WT-36 | 15-16hpf | 15/80% | 243/44.4% | 33/60.6% | 8.6 |
| WT-37 | 15-16hpf | 15/93.3% | 310/66.1% | 47/29.8% | 3.1 |

B

## hs>*cls*$^{m618}$ injections (50pg) with 2 heatshocks:

| hs>*cls*$^{m618}$ | 1$^{st}$/2$^{nd}$ heatshock | uninjected/ % survived | injected / % survived | *cls* / % rescued | mean number of M |
|---|---|---|---|---|---|
| M-1 | 9hpf/20hpf | 3/0% | 80/1% | 0/0% | - |
| M-2* | 13hpf/24hpf | 42/95% | 240/94% | 41/0% | - |
| M-3 | 12.5hpf/24.5hpf | 17/53% | 148/33% | 9/0% | - |
| M-4 | 12hpf/24hpf | 20/80% | 193/77% | 32/0% | - |
| M-5 | 12hpf/24hpf | 19/15.8% | 208/9.1% | 6/0% | - |
| M-6 | 11hpf/23hpf | 24/20.8% | 151/4.6% | 2/0% | - |
| M-7 | 10.5hpf/24hpf | 29/93% | 246/43.5% | 32/0% | - |

* The only batch that was injected into ~8 cell stage embryos. All the other hs>*cls*$^{m618}$ injections were carried out in 1-4 cell stage

## hs>*cls*$^{m618}$ injections (120pg) and one early heatshock:

| hs>*cls*$^{m618}$ | heatshock | uninjected/ % survived | injected / % survived | *cls* / % rescued | mean number of M |
|---|---|---|---|---|---|
| M-8 | 15-16hpf | 15/80% | 121/1.6% | 0/0% | nd |
| M-9 | 15-16hpf | 15/60% | 99/4% | 0/0% | nd |
| M-10 | 15-16hpf | 10/60% | 142/7% | 3/0% | - |

## hs>*cls*$^{m618}$ injections (50pg) and one early heatshock:

| hs>*cls*$^{m618}$ | heatshock | uninjected/ % survived | injected / % survived | *cls* / % rescued | mean number of M |
|---|---|---|---|---|---|
| M-11 | 15-16hpf | 15/100% | 218/3.2% | 0/0% | nd |

## hs>*cls*$^{tw2}$ injections (120pg) with one early heatshock:

| hs>*cls*$^{tw2}$ | heatshock | uninjected/ % survived | injected / % survived | *cls* / % rescued | mean number of M |
|---|---|---|---|---|---|
| T-1 | 15-16hpf | 15/100% | 226/40.3% | 18/11.1% | 3.0 |
| T-2 | 15-16hpf | 11/63.6% | 114/26.3% | 7/14.3% | 12.0 |
| T-3 | 15-16hpf | 10/20% | 209/28.2% | 2/0% | - |

## hs>*cls*$^{tw2}$ injections (50pg) with one early heatshock:

| hs>*cls*$^{tw2}$ | heatshock | uninjected/ % survived | injected / % survived | *cls* / % rescued | mean number of M |
|---|---|---|---|---|---|
| T-4 | 15-16hpf | 15/86.7% | 165/13.9% | 7/0% | - |
| T-5 | 15-16hpf | 15/86.7% | 174/61.5% | 25/0% | - |
| T-6 | 15-16hpf | 15/86.7% | 277/65.3% | 51/0% | - |
| T-7 | 15-16hpf | 15/93.3% | 191/49.7% | 23/0% | - |
| T-8 | 15-16hpf | 16/68.8% | 170/37.6% | 19/5.3% | 1.0 |
| T-9 | 15-16hpf | 15/53.3% | 151/25.8% | 7/14.3% | 2.0 |
| T-10 | 15-16hpf | 12/75% | 112/26.8% | 8/12.5% | 9.0 |
| T-15 | 15-16hpf | 11/73.3% | 80/41.9% | 19/0% | - |
| T-16 | 15-16hpf | 11/73.3% | 53/27.7% | 10/0% | - |
| T-17 | 15-16hpf | 13/86.7% | 172/60.1% | 36/0% | - |
| T-18 | 15-16hpf | 13/86.7% | 69/38.8% | 15/0% | - |

## hs>*cls*$^{tw2}$ injections (35pg) with one early heatshock:

| hs>*cls*$^{tw2}$ | heatshock | uninjected/ % survived | injected / % survived | *cls* / % rescued | mean number of M |
|---|---|---|---|---|---|
| T-11 | 15-16hpf | 7/87.5% | 22/32.4% | 1/0% | - |
| T-12 | 15-16hpf | 11/73.3% | 56/25.6% | 0/0% | - |
| T-13 | 15-16hpf | 8/53.3% | 79/45.7% | 1/0% | - |

| *sox9a* | Uninjected: survived/total (%survived) | Injected: survived/total (%survived) | *cls* survived /*cls* rescued (% rescued) | Mean number of M |
|---|---|---|---|---|
| 9A-1 | 4/15 (26.7%) | 36/357 (10.1%) | 1/0 (0%) | - |
| 9A-2 | 23/31 (74.2%) | 111/240 (46.3%) | 31/0 (0%) | - |
| 9A-3 | 16/16 (100%) | 178/239 (74.5%) | 44/0 (0%) | - |
| **Total** | 43/62 (69.4%) | 325/836 (38.9%) | 76/0 (0%) | - |

**B**

| *sox9b* | Uninjected: survived/total (%survived) | Injected: survived/total (%survived) | *cls* survived /cls rescued (% rescued) | Mean number of M |
|---|---|---|---|---|
| 9B-1 | 6/12 (50%) | 30/278 (10.8%) | 6/0 (0%) | - |
| 9B-2 | 8/8 (100%) | 69/256 (27%) | 14/0 (0%) | - |
| 9B-3 | 8/15 (53.3%) | 164/341 (48.1%) | 38/0 (0%) | - |
| **Total** | 22/35 (62.9%) | 263/875 (30.1%) | 58/0 (0%) | - |

**Appendix 7.2:** hs>*sox9a* and hs>*sox9b* injections fail to rescue melanophores in homozygous $cls^{m618}$ mutant embryos.

(A, B) Three independent batches of embryos from a heterozygous $cls^{m618}$ cross each were injected with 50pg of either hs>*sox9a* (A, 9A-1 to 9A-3) or hs>*sox9b* (B, 9B-1 to 9B-3) and heatshocked at 15-16hpf for an hour at 37°C. The two tables list the total number of uninjected (control) and injected siblings and give the number and percentage (in brackets) of embryos that had survived after 48hours. In the fourth columns, the number of *cls* embryos that survived is put in relation to the number that were rescued. The percentage of rescued embryos is noted in brackets. The last columns list the mean number of melanophores (M) with wild-type morphology observed per embryo, whereby "-" indicates no rescued melanophores. The last row (Total) allows a comparison between survival rates of injected and uninjected control embryos. The difference corresponds to embryos dying most likely due to injection trauma. Furthermore, the total number of screened *cls* embryos (*cls* survived), rescued *cls* embryos (*cls* rescued) and the total mean of melanophore number is depicted.

| hs>*sox9a* | | phenotype | Uninjected: malf/total (% malf) | Total: malf/total (% malf) | Injected: malf/total (% malf) | Total: malf/total (% malf) |
|---|---|---|---|---|---|---|
| 9A-1 | | WT | 0/3 (0%) | 0/33 (0%) | 29/35 (82.9%) | 195/249 (78.3%) |
| | | *cls* | 0/1 (0%) | 0/10 (0%) | 0/1 (0%) | 60/76 (78.9%) |
| 9A-2 | 50pg | WT | 0/17 (0%) | | 56/80 (70%) | |
| | early hs | *cls* | 0/6 (0%) | | 24/31 (77.4%) | |
| 9A-3 | | WT | 0/13 (0%) | | 110/134 (82.1%) | |
| | | *cls* | 0/3 (0%) | | 36/44 (81.8%) | |

| hs>*sox9b* | | phenotype | Uninjected: malf/total (% malf) | Total: malf/total (% malf) | Injected: malf/total (% malf) | Total: malf/total (% malf) |
|---|---|---|---|---|---|---|
| 9B-1 | | WT | 0/6 (0%) | 0/20 (0%) | 23/24 (95.8%) | 153/205 (74.6%) |
| | | *cls* | 0/0 (0%) | 0/2 (0%) | 6/6 (100%) | 42/58 (72.4%) |
| 9B-2 | 50pg | WT | 0/7 (0%) | | 43/55 (78.2%) | |
| | early hs | *cls* | 0/1 (0%) | | 13/14 (92.9%) | |
| 9B-3 | | WT | 0/7 (0%) | | 87/126 (69%) | |
| | | *cls* | 0/1 (0%) | | 23/38 (60.5%) | |

**Appendix 7.3:** Summary of malformation data

In the first column the Table lists the names of batches of embryos and constructs injected (WT 3-37, hs>*cls*$^{WT}$; pCS 1-11, control plasmid pCS2+; M 3-11, hs>*cls*$^{m618}$; T 1-18, hs>*cls*$^{tw2}$; 9A 1-3, hs>*sox9a*; 9B 1-3, hs>*sox9b*) and indicates which were injected on the same day (brackets), the approximate amount of construct stock solution injected per embryo and the number and time of heatshock treatment(s) (early, 15-16hpf; late, 18-19.5hpf; hs, heatshock). The column labelled "phenotype" indicates whether subsequent columns refer to WT or *cls* embryos. Column 3 and 5 list the number of malformed embryos within the total number of embryos (malf/total) and the percentage of malformed embryos (% malf) in phenotypically wild-type or homozygous *cls* mutant siblings in uninjected (column 3) and injected batches (column 5). Column 4 and 6 contain the number of total malformed/total embryos (% malf) for all injected (column 6) and uninjected batches (column 4) for each injection day (brackets in column 1). nd, not determined.

| hs>*cls*$^{WT}$ | | phenotype | Uninjected: malf/total (% malf) | Total: malf/total (% malf) | Injected: malf/total (% malf) | Total: malf/total (% malf) |
|---|---|---|---|---|---|---|
| WT-3 | | WT | 0/6 (0%) | 1/12 (8%) | 23/38 (60.5%) | 82/132 (62.1%) |
| | | *cls* | 0/2 (0%) | 0/16 (0%) | 9/12 (75%) | 29/42 (69%) |
| WT-4 | 50pg | WT | nd | | 5/7 (71.4%) | |
| | 2 hs | *cls* | nd | | 3/3 (100%) | |
| WT-5 | | WT | 0/4 (0%) | | 54/87 (62%) | |
| | | *cls* | 1/10 (10%) | | 17/27 (63%) | |
| WT-6 | | WT | 0/10 (0%) | 1/54 (1.9%) | 7/27 (25.9%) | 104/186 (55.9%) |
| | | *cls* | 0/0 (0%) | 1/15 (6.7%) | 4/10 (40%) | 25/48 (52.1%) |
| WT-7 | | WT | 1/6 (16.6%) | | 17/25 (68%) | |
| | | *cls* | 1/2 (50%) | | 6/7 (85.7%) | |
| WT-8 | | WT | 0/0 (0%) | | 9/13 (69.2%) | |
| | 50pg | *cls* | 0/0 (0%) | | 0/1 (0%) | |
| WT-9 | no hs | WT | 0/11 (0%) | | 16/25 (64%) | |
| | | *cls* | 0/3 (0%) | | 3/4 (75%) | |
| WT-10 | | WT | 0/22 (0%) | | 47/64 (73.4%) | |
| | | *cls* | 0/7 (0%) | | 8/14 (57.1%) | |
| WT-11 | | WT | 0/5 (0%) | | 8/32 (25%) | |
| | | *cls* | 0/3 (0%) | | 4/12 (33%) | |
| WT-12 | | WT | 0/0 (0%) | 1/52 (1.9%) | 3/3 (100%) | 30/169 (17.8%) |
| | 25pg | *cls* | 1/2 (50%) | 1/23 (4.3%) | 0/0 (0%) | 12/48 (25%) |
| WT-13 | early hs | WT | 0/7 (0%) | | 17/57 (29.8%) | |
| | | *cls* | 0/2 (0%) | | 6/21 (28.6%) | |
| WT-14 | 25pg | WT | 0/12 (0%) | 0/12 (0%) | 37/165 (22.4%) | 37/165 (22.4%) |
| | late hs | *cls* | 0/6 (0%) | 0/6 (0%) | 8/51 (15.7%) | 8/51 (15.7%) |
| WT-15 | | WT | 0/15 (0%) | see WT-12 | 9/78 (11.5%) | see WT-12 |
| | 25pg | *cls* | 0/5 (0%) | | 5/19 (26.3%) | |
| WT-16 | early hs | WT | 1/10 (10%) | | 1/31 (3.2%) | |
| | | *cls* | 0/3 (0%) | | 1/8 (12.5%) | |
| WT-17 | | WT | 0/10 (0%) | 0/20 (0%) | 11/30 (36.7%) | 37/100 (37%) |
| | 50pg | *cls* | 0/5 (0%) | 0/11 (0%) | 5/19 (55.5%) | 14/43 (32.6%) |
| WT-18 | early hs | WT | 0/10 (0%) | | 26/70 (37.1%) | |
| | | *cls* | 0/6 (0%) | | 9/24 (37.5%) | |
| WT-19 | | WT | nd | 0/33 (0%) | 0/0 (0%) | 141/152 (92.8%) |
| | | *cls* | nd | 0/15 (0%) | 0/0 (0%) | 36/43 (83.7%) |
| WT-20 | | WT | 0/11 (0%) | | 39/41 (95.1%) | |
| | | *cls* | 0/5 (0%) | | 13/13 (100%) | |
| WT-21 | | WT | 0/5 (0%) | | 29/29 (100%) | |
| | 50pg | *cls* | 0/3 (0%) | | 6/6 (100%) | |
| WT-22 | late hs | WT | 0/8 (0%) | | 39/46 (84.8%) | |
| | | *cls* | 0/2 (0%) | | 9/16 (56.3%) | |
| WT-23 | | WT | 0/3 (0%) | | 14/14 (100%) | |
| | | *cls* | 0/2 (0%) | | 1/1 (100%) | |
| WT-24 | | WT | 0/6 (0%) | | 20/22 (90.9%) | |
| | | *cls* | 0/3 (0%) | | 7/7 (100%) | |
| WT-25 | | WT | 0/9 (0%) | 1/124 (0.8%) | 10/59 (16.9%) | 51/360 (14.2%) |
| | | *cls* | 0/1 (0%) | 0/20 (0%) | 5/28 (17.8%) | 21/122 (17.2%) |
| WT-26 | | WT | 0/6 (0%) | | 9/25 (36%) | |
| | | *cls* | 0/1 (0%) | | 2/8 (25%) | |
| WT-27 | 25pg | WT | 0/13 (0%) | | 23/120 (19.2%) | |
| | no hs | *cls* | 0/3 (0%) | | 12/43 (27.9%) | |
| WT-28 | | WT | 1/96 (1%) | | 8/137 (5.8%) | |
| | | *cls* | 0/15 (0%) | | 1/36 (2.8%) | |
| WT-29 | | WT | nd | | 1/19 (5.3%) | |
| | | *cls* | nd | | 1/7 (14.3%) | |

| | | | | | | |
|---|---|---|---|---|---|---|
| WT-30 | | WT | 3/8 (37.5%) | 4/26 (15.4%) | 11/23 (47.8%) | 94/236 (39.8%) |
| | | *cls* | 1/2 (50%) | 1/5 (20.0%) | 4/5 (80%) | 33/82 (40.2%) |
| WT-31 | | WT | 0/0 (0%) | | 9/18 (50%) | |
| | | *cls* | 0/0 (0%) | | 6/9 (66.6%) | |
| WT-32 | | WT | 0/4 (0%) | | 13/23 (56.5%) | |
| | | *cls* | 0/1 (0%) | | 7/14 (50%) | |
| WT-33 | 12pg | WT | nd | | 2/3 (75%) | |
| | early hs | *cls* | nd | | 0/1 (0%) | |
| WT-34 | | WT | 1/2 (50%) | | 4/4 (100%) | |
| | | *cls* | 0/0 (0%) | | 0/0 (0%) | |
| WT-38 | | WT | nd | | 14/75 (18.7%) | |
| | | *cls* | nd | | 5/20 (25%) | |
| WT-39 | | WT | 0/12 (0%) | | 41/90 (45.6%) | |
| | | *cls* | 0/2 (0%) | | 11/33 (33.3%) | |
| WT-35 | | WT | 0/5 (0%) | 0/22 (0%) | 42/62 (67.7%) | 216/294 (73.5%) |
| | | *cls* | 0/3 (0%) | 0/12 (0%) | 12/13 (92.3%) | 70/93 (75.3%) |
| WT-36 | 120pg | WT | 0/8 (0%) | | 56/74 (75.7%) | |
| | early hs | *cls* | 0/4 (0%) | | 26/33 (78.8%) | |
| WT-37 | | WT | 0/9 (0%) | | 118/158 (74.7%) | |
| | | *cls* | 0/5 (0%) | | 32/47 (68.1%) | |

| **pCS only** | | **phenotype** | **Uninjected: malf/total (% malf)** | **Total: malf/total (% malf)** | **Injected: malf/total (% malf)** | **Total: malf/total (% malf)** |
|---|---|---|---|---|---|---|
| pCS-1 | | WT | 0/12 (0%) | 0/12 (0%) | 25/41 (61%) | 87/146 (59.6%) |
| | | *cls* | 0/2 (0%) | 0/2 (0%) | 11/20 (55%) | 23/46 (50%) |
| pCS-3 | 50pg | WT | nd | | 11/18 (61.1%) | |
| | early hs | *cls* | nd | | 1/3 (33.3%) | |
| pCS-4 | | WT | nd | | 51/87 (58.6%) | |
| | | *cls* | nd | | 11/23 (47.8%) | |
| pCS-2 | | WT | 0/4 (0%) | 0/13 (0%) | 35/41 (85.4%) | 102/137 (74.5%) |
| | 50pg | *cls* | 0/5 (0%) | 0/7 (0%) | 14/15 (93.3%) | 27/37 (73%) |
| pCS-5 | late hs | WT | 0/9 (0%) | | 67/96 (69.8%) | |
| | | *cls* | 0/2 (0%) | | 13/22 (59.1%) | |
| pCS-6 | | WT | 0/2 (0%) | 0/28 (0%) | 1/2 (50%) | 57/146 (39%) |
| | | *cls* | 0/0 (0%) | 0/8 (0%) | 0/0 (0%) | 18/51 (35.3%) |
| pCS-7 | | WT | 0/6 (0%) | | 12/25 (48%) | |
| | 25pg | *cls* | 0/2 (0%) | | 6/11 (54.5%) | |
| pCS-8 | late hs | WT | 0/10 (0%) | | 35/75 (46.7%) | |
| | | *cls* | 0/3 (0%) | | 10/28 (35.7%) | |
| pCS-9 | | WT | 0/10 (0%) | | 9/44 (20.5%) | |
| | | *cls* | 0/3 (0%) | | 2/12 (16.7%) | |
| pCS-10 | | WT | 0/7 (0%) | 0/7 (0%) | 15/43 (34.9%) | 38/118 (32.2%) |
| | 25pg | *cls* | 0/7 (0%) | 0/7 (0%) | 2/9 (22.2%) | 6/27 (22.2%) |
| pCS-11 | early hs | WT | nd | | 23/75 (30.7%) | |
| | | *cls* | nd | | 4/18 (22.2%) | |

| hs>*cls*$^{m618}$ | | phenotype | Uninjected: malf/total (% malf) | Total: malf/total (% malf) | Injected: malf/total (% malf) | Total: malf/total (% malf) |
|---|---|---|---|---|---|---|
| M-3 | | WT | 0/4 (0%) | 2/39 (5.1%) | 7/35 (20%) | 72/231 (31.2%) |
| | | *cls* | 0/4 (0%) | 0/20 (0%) | 1/9 (11%) | 23/81 (28.4%) |
| M-4 | | WT | 0/12 (0%) | | 21/103 (20%) | |
| | | *cls* | 0/4 (0%) | | 6/32 (18.8%) | |
| M-5 | 50pg | WT | 0/3 (0%) | | 6/13 (46%) | |
| | 2hs | *cls* | 0/0 (0%) | | 2/6 (33%) | |
| M-6 | | WT | 0/3 (0%) | | 3/5 (60%) | |
| | | *cls* | 0/2 (0%) | | 1/2 (50%) | |
| M-7 | | WT | 2/17 (11.8%) | | 35/75 (46.6%) | |
| | | *cls* | 0/10 (0%) | | 13/32 (40.6%) | |
| M-8 | | WT | 0/7 (0%) | 0/20 (0%) | 2/2 (100%) | 13/13 (100%) |
| | | *cls* | 0/5 (0%) | 0/7 (0%) | 0/0 (/) | 3/3 (100%) |
| M-9 | 120pg | WT | 0/7 (0%) | | 4/4 (100%) | |
| | early hs | *cls* | 0/2 (0%) | | 0/0 (/) | |
| M-10 | | WT | 0/6 (0%) | | 7/7 (100%) | |
| | | *cls* | 0/0 (0%) | | 3/3 (100%) | |
| M-11 | 50pg | WT | 0/11 (0%) | 0/11(0%) | 7/7 (100%) | 7/7 (100%) |
| | early hs | *cls* | 1/4 (25%) | 1/4 (25%) | nd | nd |

| hs>*cls*$^{tw2}$ | | phenotype | Uninjected: malf/total (% malf) | Total: malf/total (% malf) | Injected: malf/total (% malf) | Total: malf/total (% malf) |
|---|---|---|---|---|---|---|
| T-1 | | WT | 0/9 (0%) | 0/17 (0%) | 66/73 (90.4%) | 138/153 (90.2%) |
| | | *cls* | 0/6 (0%) | 0/7 (0%) | 17/18 (94.4%) | 24/27 (88.9%) |
| T-2 | 120pg | WT | 0/6 (0%) | | 21/23 (91.3%) | |
| | early hs | *cls* | 0/1 (0%) | | 6/7 (85.7%) | |
| T-3 | | WT | 0/2 (0%) | | 51/57 (89.5%) | |
| | | *cls* | 0/0 (0%) | | 1/2 (50%) | |
| T-4 | | WT | 0/7 (0%) | 0/35 (0%) | 7/16 (43.8%) | 170/300 (56.7%) |
| | | *cls* | 0/6 (0%) | 0/18 (0%) | 3/7 (42.9%) | 47/106 (44.3%) |
| T-5 | | WT | 0/10 (0%) | | 45/82 (54.9%) | |
| | 50pg | *cls* | 0/3 (0%) | | 14/25 (56%) | |
| T-6 | early hs | WT | 0/9 (0%) | | 72/130 (55.4%) | |
| | | *cls* | 0/4 (0%) | | 21/51 (41.2%) | |
| T-7 | | WT | 0/9 (0%) | | 46/72 (63.9%) | |
| | | *cls* | 0/5 (0%) | | 9/23 (39.1%) | |
| T-8 | | WT | 1/6 (16.7%) | 2/20 (10%) | 43/45 (95.6%) | 97/99 (98%) |
| | | *cls* | 1/4 (25%) | 1/7 (14.3%) | 16/19 (84.2%) | 31/34 (91.2%) |
| T-9 | 50pg | WT | 0/8 (0%) | | 32/32 (100%) | |
| | early hs | *cls* | nd | | 7/7 (100%) | |
| T-10 | | WT | 1/6 (16.7%) | | 22/22 (100%) | |
| | | *cls* | 0/3 (0%) | | 8/8 (100%) | |
| T-11 | | WT | 0/7 (0%) | 0/21 (0%) | 11/11 (100%) | 17/18 (94.4%) |
| | | *cls* | 0/0 (0%) | 0/5 (0%) | 1/1 (100%) | 2/2 (100%) |
| T-12 | 35pg | WT | 0/8 (0%) | | 3/4 (75%) | |
| | early hs | *cls* | 0/3 (0%) | | 0/0 (/%) | |
| T-13 | | WT | 0/6 (0%) | | 3/3 (100%) | |
| | | *cls* | 0/2 (0%) | | 1/1 (100%) | |
| T-15 | | WT | 0/7 (0%) | 1/58 (1.7%) | 43/61 (70.5%) | 194/294 (66%) |
| | | *cls* | 0/4 (0%) | 0/17 (0%) | 12/19 (63.2%) | 50/80 (62.5%) |
| T-16 | | WT | 0/19 (0%) | | 29/43 (67.4%) | |
| | 50pg | *cls* | 0/2 (0%) | | 8/10 (80%) | |
| T-17 | early hs | WT | 1/9 (11.1%) | | 91/136 (66.9%) | |
| | | *cls* | 0/4 (0%) | | 23/36 (63.9%) | |
| T-18 | | WT | 0/6 (0%) | | 31/54 (57.4%) | |
| | | *cls* | 0/7 (0%) | | 7/15 (46.7%) | |

**Appendix 7.4:** Statistical test results assessing various influences on malformation (A) With this $\chi^2$ test it was asked whether *cls* mutant embryos were affected more severely by malformations and thus, whether the genotype had any influence on the occurrence of malformations. Batches of embryos injected with hs>$cls^{WT}$ (WT3-29), control plasmid (pCS1-11) or hs>$cls^{m618}$ (M3-7) were assessed in groups corresponding to groups with different treatments (heatshock times, construct concentrations etc, Appendix 7.1). The column "WT ratio" lists the ratio of malformed:normal embryos with a wild-type phenotype (WT). "*cls* observed" contains the number of malformed homozygous *cls* mutants, whereas "*cls* expected" is the number of malformed *cls* embryos that would be expected to be deformed, if the same ratio applied to *cls* as did for WT. The $\chi^2$-test assesses whether there is a significant difference between the number of "*cls* observed" and "*cls* expected". If the calculated $\chi^2$ value in column "$\chi^2$" is larger than the tabulated one at a particular confidence level ($\chi^2_{df,\ 5\text{-}0.1\%}$) there is a significant difference (5% level) or even highly significant difference (0.1% level) between the expected and observed numbers of malformed *cls* embryos. (B-E) Influences on malformation frequency was tested with a single factor ANOVAR analysis, which tests whether the variation between different treatments is any different to the variation observed within injection batches of the same treatment. Columns labelled hs (heatshock) indicate whether an early (15-16hpf), late (18-19.5hpf) or no heatshock (no hs) had been carried out. The column "DNA amount" shows whether a high (50pg) or low (25pg) amount of a construct (hs>$cls^{WT}$ or pCS) had been injected. If the calculated F value (F) is higher than the tabulated F value ($F_{df1,\ df2,\ \%}$) at a particular confidence level (0.1%, 1% or 5%), there is a significant difference (5% level) or even a highly significant difference (0.1% level) between variation within and between groups of treatments. df, degree of freedom. (F) A $\chi$2-test was used to test whether the ratio of the number of rescued embryos with normal morphology and malformed rescued embryos was different to a ratio of 1:1. It was asked whether malformed embryos were less likely rescued and thus, whether the presence of malformations had any influence on the rescue. The column "treatment" describes the DNA concentration, number and time of heatshock as outlined in (B-E). The column "malf.:normal" lists the ratio of malformed (malf.) to normal rescued *cls* embryos taken from Appendix 7.1. If the calculated $\chi$2-value is smaller than the tabulated one ($\chi^2_{df,\ 5\%}$) for the 5% confidence level, there is no significant difference from the ratio 1:1.

## A) Do malformations vary with the genotype (WT vs. *cls*$^{-/-}$)?

$\chi^2$-test on *cls*/WT pairs of individual datasets using the Yates' correction formula.
Degree of freedom =1 in all datasets: $\chi^2_{1,5\%} = 3.841$, $\chi^2_{1,1\%} = 6.635$, $\chi^2_{1,0.1\%} = 7.879$

| | WT ratio | *cls* observed | *cls* expected | $\chi^2$ | Result |
|---|---|---|---|---|---|
| WT-3 | 23/38 | 9 | 7.3 | 0.2788 | no sign diff on any level |
| WT-4 | 5/7 | 3 | 2.1 | | |
| WT-5 | 54/87 | 17 | 16.8 | | |
| WT-6 | 7/27 | 4 | 2.6 | 0.8725 | no sign diff on any level |
| WT-7 | 17/25 | 6 | 4.8 | | |
| WT-8 | 9/13 | 0 | 0.7 | | |
| WT-9 | 16/25 | 3 | 2.6 | | |
| WT-10 | 47/64 | 8 | 10.3 | | |
| WT-11 | 8/32 | 4 | 3.0 | | |
| WT-13 | 17/57 | 6 | 6.3 | 2.544 | no sign diff on any level |
| WT-15 | 9/78 | 5 | 2.2 | | |
| WT-16 | 1/31 | 1 | 0.3 | | |
| WT-14 | 37/165 | 8 | 11.4 | 0.7377 | no sign diff on any level |
| WT-17 | 11/30 | 5 | 7.0 | 0.3394 | no sign diff on any level |
| WT-18 | 26/70 | 9 | 8.9 | | |
| WT-20 | 39/41 | 13 | 12.4 | 1.5301 | no sign diff on any level |
| WT-21 | 29/29 | 6 | 6.0 | | |
| WT-22 | 39/46 | 9 | 13.6 | | |
| WT-23 | 14/14 | 1 | 1.0 | | |
| WT-24 | 20/22 | 7 | 6.4 | | |
| WT-25 | 10/59 | 5 | 4.7 | 1.6203 | no sign diff on any level |
| WT-26 | 9/25 | 2 | 2.9 | | |
| WT-27 | 23/120 | 12 | 8.2 | | |
| WT-28 | 8/137 | 1 | 2.1 | | |
| WT-29 | 1/19 | 1 | 0.7 | | |
| pCS-1 | 25/41 | 11 | 12.2 | 0.3865 | no sign diff on any level |
| pCS-3 | 11/18 | 1 | 1.8 | | |
| pCS-4 | 51/87 | 11 | 13.5 | | |
| pCS-2 | 35/41 | 14 | 12.8 | 0.2727 | no sign diff on any level |
| pCS-5 | 67/96 | 13 | 15.4 | | |
| pCS-7 | 12/25 | 6 | 5.3 | 0.5236 | no sign diff on any level |
| pCS-8 | 35/75 | 10 | 13.1 | | |
| pCS-9 | 9/44 | 2 | 2.5 | | |
| pCS-10 | 15/43 | 2 | 3.1 | 0.2979 | no sign diff on any level |
| pCS-11 | 23/75 | 4 | 5.5 | | |
| M-3 | 7/35 | 1 | 1.8 | 0.2887 | no sign diff on any level |
| M-4 | 21/103 | 6 | 6.5 | | |
| M-5 | 6/13 | 2 | 2.8 | | |
| M-6 | 3/5 | 1 | 1.2 | | |
| M-7 | 35/75 | 13 | 14.9 | | |

## B) Does a heatshock increase malformations?

Single factor anovar analysis with 2 levels

| hs | DNA amount | construct | F | $F_{df1, df2, \%}$ | Result |
|---|---|---|---|---|---|
| no hs | high | hs>*cls*$^{WT}$ | 92.88 | 0.1% - 17.82 | sign. diff. even on 0.1% level |
| early | high | hs>*cls*$^{WT}$ | | | |
| no hs | high | hs>*cls*$^{WT}$ | 21.24 | 0.1% - 15.08 | sign. diff. even on 0.1% level |
| late | high | hs>*cls*$^{WT}$ | | | |
| no hs | low | hs>*cls*$^{WT}$ | 3.24 | 5% - 4.75 | no sign. diff. even on 5% level |
| early | low | hs>*cls*$^{WT}$ | | | |
| no hs | low | hs>*cls*$^{WT}$ | 0.057 | 5% - 4.96 | no sign. diff. even on 5% level |
| late | low | hs>*cls*$^{WT}$ | | | |

## C) Is there a difference in malformation between hs>clsWT and pCS?

Single factor anovar analysis with 2 levels

| construct | DNA amount | hs | F | $F_{df1, df2, \%}$ | Result |
|---|---|---|---|---|---|
| hs>*cls*$^{WT}$ | high | early | 9.81 | 5% - 5.32 | sign. diff. only on 5% level |
| pCS | high | early | | | |
| hs>*cls*$^{WT}$ | high | late | 3.52 | 5% - 4.75 | no sign. diff. even on 5% level |
| pCS | high | late | | | |
| hs>*cls*$^{WT}$ | low | early | 0.027 | 5% - 5.12 | no sign. diff. even on 5% level |
| pCS | low | early | | | |
| hs>*cls*$^{WT}$ | low | late | 3.10 | 5% - 5.59 | no sign. diff. even on 5% level |
| pCS | low | late | | | |

## D) Is there a difference in malformation between high and low DNA concentration?

Single factor anovar analysis with 2 levels

| DNA amount | construct | hs | F | $F_{df1, df2, \%}$ | Result |
|---|---|---|---|---|---|
| high | hs>*cls*$^{WT}$ | early | 0.059 | 5% - 5.12 | no sign. diff. even on 5% level |
| low | hs>*cls*$^{WT}$ | early | | | |
| high | hs>*cls*$^{WT}$ | late | 51.84 | 0.1% - 21.04 | sign. diff. even on 0.1% level |
| low | hs>*cls*$^{WT}$ | late | | | |
| high | pCS | early | 17.70 | 1% - 11.26 | sign. diff. on 1% level |
| low | pCS | early | | | |
| high | pCS | late | 16.04 | 1% - 10.56 | sign. diff. on 1% level |
| low | pCS | late | | | |

## E) Do malformations change depending on when the heatshock was carried out?

Single factor anovar analysis with 2 levels

| hs time | DNA amount | construct | F | $F_{df1, df2, \%}$ | Result |
|---|---|---|---|---|---|
| early | high | hs>*cls*$^{WT}$ | 64.36 | 0.1% - 18.64 | sign. diff. even on 0.1% level |
| late | high | hs>*cls*$^{WT}$ | | | |
| early | high | pCS | 8.67 | 5% - 5.32 | sign. diff. only on 5% level |
| late | high | pCS | | | |
| early | low | hs>*cls*$^{WT}$ | 0.217 | 5% - 5.59 | no sign. diff. even on 5% level |
| late | low | hs>*cls*$^{WT}$ | | | |
| early | low | pCS | 2.0 | 5% - 5.12 | no sign. diff. even on 5% level |
| late | low | pCS | | | |

## F) Are malformed embryos more likely to be rescued? Hence, does malformation have any effect on the rescue experiment?

$\chi^2$-test to compare whether there is a significant difference between the ratio normal:malformed and 1:1 within the population of rescued embryos.
$H_0$: normal/malf = 1:1
$\chi^2$-tests carried out on all 22 datasets and also within groups of different treatments.

| | treatment | normal/malf | $\chi^2$ (1:1) | $\chi^2_{df, 5\%}$ | Result |
|---|---|---|---|---|---|
| WT-G | high, 2 hs | 8:9 | 2.778 | 7.815 | no sign diff |
| WT-3 | | 0:5 | | | |
| WT-4 | | 0:3 | | | |
| WT-5 | | 1:8 | | | |
| WT-7 | high, no hs | 0:2 | 2.111 | 7.815 | no sign diff |
| WT-9 | | 0:3 | | | |
| WT-10 | | 4:3 | | | |
| WT-11 | | 2:2 | | | |
| WT-13 | low, early | 4:1 | 1.5625 | 5.991 | no sign diff |
| WT-15 | | 3:2 | | | |
| WT-16 | | 1:0 | | | |
| WT-14 | low, late | 11:5 | na | na | na |
| WT-17 | high, early | 4:3 | 0.1389 | 3.841 | no sign diff |
| WT-18 | | 7:6 | | | |
| WT-20 | high, late | 1:7 | 3.179 | 7.815 | no sign diff |
| WT-21 | | 0:3 | | | |
| WT-22 | | 1:3 | | | |
| WT-24 | | 0:6 | | | |
| WT-25 | | 2:1 | 2.25 | 7.815 | no sign diff |
| WT-26 | | 1:0 | | | |
| WT-27 | | 4:2 | | | |
| WT-28 | | 2:0 | | | |

Development 128, 4113-4125 (2001)
Printed in Great Britain © The Company of Biologists Limited 2001
DEV1668



# Zebrafish *colourless* encodes *sox10* and specifies non-ectomesenchymal neural crest fates

**Kirsten A. Dutton[1,*], Angela Pauliny[1,*], Susana S. Lopes[1], Stone Elworthy[1], Tom J. Carney[1], Jörg Rauch[2], Robert Geisler[2], Pascal Haffter[2] and Robert N. Kelsh[1,‡]**

[1]Department of Biology and Biochemistry, University of Bath, Claverton Down, Bath BA2 7AY, UK
[2]Max-Planck-Institut für Entwicklungsbiologie, Spemannstraße 35/III, D-72076 Tübingen, Germany
*These authors contributed equally to this work and are to be considered joint first authors
‡Author for correspondence (e-mail: bssrnk@bath.ac.uk)

*Accepted 24 July 2001*

## SUMMARY

**Waardenburg-Shah syndrome combines the reduced enteric nervous system characteristic of Hirschsprung's disease with reduced pigment cell number, although the cell biological basis of the disease is unclear. We have analysed a zebrafish Waardenburg-Shah syndrome model. We show that the *colourless* gene encodes a *sox10* homologue, identify *sox10* lesions in mutant alleles and rescue the mutant phenotype by ectopic *sox10* expression. Using iontophoretic labelling of neural crest cells, we demonstrate that *colourless* mutant neural crest cells form ectomesenchymal fates. By contrast, neural crest cells which in wild types form non-ectomesenchymal fates generally fail to migrate and do not overtly differentiate. These cells die by apoptosis between 35 and 45 hours post fertilisation. We provide evidence that melanophore defects in *colourless* mutants can be largely explained by disruption of *nacre/mitf* expression. We propose that all defects of affected crest derivatives are consistent with a primary role for *colourless/sox10* in specification of non-ectomesenchymal crest derivatives. This suggests a novel mechanism for the aetiology of Waardenburg-Shah syndrome in which affected neural crest derivatives fail to be generated from the neural crest.**

Key words: *Danio rerio*, Waardenburg-Shah syndrome, Hirschsprung's disease, Pigment cells, Melanophore, Apoptosis

## INTRODUCTION

The neural crest is a vertebrate tissue of developmental and medical importance. Developmentally, neural crest is intriguing because the cells are initially multipotent and subsequently form a great diversity of derivative cell types, including ectomesenchymal fates, such as craniofacial skeleton and fin mesenchyme, and non-ectomesenchymal fates, such as neurones, glia and pigment cells (Le Douarin, 1982; Smith et al., 1994). Medically, neural crest is important because some diseases, known as neurocristopathies and including diverse conditions such as albinism, neurofibromatosis and Hirschsprung's disease (aganglionic megacolon), affect cell types derived from this tissue (Bolande, 1974).

Understanding the genetic and embryological basis of neurocristopathies has depended on animal models. Thus, models for Hirschsprung's disease, in which individuals have few or no enteric ganglia in the colon, or the related Waardenburg-Shah syndrome, in which individuals combine Hirschsprung's disease with pigmentary anomalies of the skin, hair and irises, have been described in several species, including mouse and zebrafish (Hosoda et al., 1994; Kelsh and Eisen, 2000). Analysis of such models in mice has identified three loci that are crucial for Waardenburg-Shah syndrome (Attié et al., 1995; Edery et al., 1996; Hofstra et al., 1996; Pingault et al., 1998; Puffenberger et al., 1994; Southard-Smith et al., 1999). Thus, mutations in loci encoding the G-protein-coupled transmembrane receptor protein endothelin receptor B (Ednrb) or its natural ligand endothelin 3 (Edn3) result in aganglionosis of terminal gut in homozygous mutants (Baynash et al., 1994; Hosoda et al., 1994), as do heterozygous mutations in the *Sry*-related transcription factor gene *Sox10* (Herbarth et al., 1998; Southard-Smith et al., 1998). Homozygous *Sox10* mutant animals show a more severe phenotype with aganglionosis of the whole gut (Herbarth et al., 1998; Southard-Smith et al., 1998). Additionally, mutations in all these genes affect body pigmentation (Lane and Liu, 1984; Mayer, 1965; Mayer and Maltby, 1964), but only *Sox10* mutations result in widespread peripheral nervous system defects (Herbarth et al., 1998; Kapur, 1999; Southard-Smith et al., 1998).

The Sox gene family encodes a large family of transcription factors, with vertebrates likely to have more than 20 Sox genes each (Wegner, 1999). Their precise roles are not well understood, although many are presumed to function in cell fate specification (Pevny and Lovell-Badge, 1997). For example, the founding family member, *Sry*, is likely to be responsible for Sertoli cell specification, and thus male sex

determination, in mammals. While *Sox10* is clearly an important transcriptional regulator in neural crest cell (NCC) development, the cellular basis of the *Sox10* mutant phenotype remains unclear. It has been suggested that peripheral nervous system and pigmentation defects result from loss of NCCs (Southard-Smith et al., 1998; Kapur 1999), although the developmental status of these cells at the time of loss is unknown. Furthermore, roles in defining regional identity in the cranial neural crest and in glial cell differentiation have also been proposed (Bondurand et al., 1998; Herbarth et al., 1998; Kuhlbrodt et al., 1998a; Pusch et al., 1998; Southard-Smith et al., 1998; Britsch et al., 2001).

Mutations at the *colourless* (*cls*) locus have been identified in zebrafish mutagenesis screens (Kelsh et al., 1996; Malicki et al., 1996). We have previously characterised the crest derivative defect displayed by *cls* mutants (Kelsh et al., 2000a; Kelsh and Eisen, 2000), noting extensive loss of pigment cells and enteric nervous system, together with large reductions in sensory and sympathetic neurones and putative satellite glia and Schwann cells. By contrast, we found little effect on ectomesenchymal derivatives, craniofacial skeleton and fin mesenchyme. Based on the severity and details of the phenotype, we proposed that *cls* functions in specification, proliferation or survival of a progenitor(s) for all non-ectomesenchymal crest derivatives.

The *cls* phenotype, and the cell-autonomy of *cls* gene action in pigment cell types (Kelsh and Eisen, 2000), suggested *sox10* as a candidate gene. We provide an experimental test of this hypothesis. We report the mapping of the *cls* locus and cloning of a zebrafish *sox10* homologue. We show linkage between *cls* and *sox10*, identify *sox10* lesions in four mutant alleles and show rescue of the *cls* phenotype by *sox10* expression. In addition, we describe iontophoretic labelling experiments to examine the precise cell-biological role of *cls/sox10* gene function in neural crest development. We show in live embryos that NCC clones in *cls* and wild-type embryos differentiated into ectomesenchymal fates after migration to appropriate sites. Remaining NCC clones adopted non-ectomesenchymal fates in wild-type embryos. By contrast, in *cls* embryos differentiation to non-ectomesenchymal fates was rarely observed. Instead, most clones failed to migrate and underwent late cell death by an apoptotic mechanism. Finally, for the melanophore fate, we show disrupted expression in *cls* mutants of genes vital for melanophore specification and migration. Together, these data demonstrate a complex phenotype in *cls* embryos that can be explained by proposing that *cls/sox10* has a primary role in specification of non-ectomesenchymal fates. Defects in cell migration, survival and differentiation are therefore likely to be secondary consequences of an inability of these cells to adopt specific fates.

## MATERIALS AND METHODS

### Fish husbandry

Embryos were obtained through natural crosses and staged according to Kimmel et al. (Kimmel et al., 1995). We used 4 *cls* alleles (*m618*, *t3*, *tw2* and *tw11*).

### Mapping of *cls*

$cls^{tw11}$ heterozygous fish (Tübingen background), were crossed to wild-type strain WIK11 to produce a reference mapping cross. Heterozygous $F_1$ were incrossed and separate pools of 12 homozygous *cls* fish and their wild-type siblings were used for simple sequence length polymorphism analysis (Knapik et al., 1996). Linkages from the pools were confirmed and refined by genotyping individual embryos, as described by Rauch et al. (Rauch et al., 1997).

### Isolation, sequencing, phylogenetic analysis and radiation hybrid mapping of zebrafish *sox10*

RT-PCR was performed using total RNA of 19 hpf stage wild-type embryos using published conditions and degenerate primers (Yuan et al., 1995). Sequencing of resulting clones identified a *sox10*-like sequence. The *sox10* clone was extended by RACE PCR using gene-specific primers (Clontech, SMART RACE kit) and sequenced on an ABI DNA sequencer. All primer sequences available on request. The full *sox10* cDNA sequence is available in Genbank (Accession Number AF402677. The zebrafish *sox10* homologue was mapped on the radiation hybrid panel LN54 (Hukriede et al., 1999) by PCR with primers 5′-ACCGTGACACACTCTACCAAGATGACC-3′ and 5′-CATGATAAAATTTGCACCCTGAAAAGG-3′, which generate a 931bp 3′ UTR fragment.

For phylogenetic analysis sequences were extracted from Genbank and coding sequences automatically extracted using Genetrans (within GCG9). These were translated and then aligned using ClustalX (Thompson et al., 1997). The nucleotide alignments were reconstructed from the protein alignments using MRTRANS (www.hgmp.mrc.ac.uk/Registered/Option/mrtrans.html). Tree-Puzzle (v 4.0.2) was used to construct an unrooted tree by maximum likelihood (Strimmer and von Haeseler, 1996). It automatically assigns estimations of support to each internal branch, figures for which are presented in the figure. To model the substitution process the Tamura and Nei (Tamura and Nei, 1993) model was employed and all sites were used. Gamma distributed variation in rates of evolution was permitted with eight variable sites and one invariable. Parameters were estimated using quartet sampling and the neighbour-joining tree.

### Characterisation of mutant *sox10* alleles

Total RNA from 27 hpf homozygous embryos of mutant alleles $cls^{m618}$, $cls^{tw2}$ and $cls^{tw11}$ was prepared using TRI reagent (Sigma). First strand cDNA was generated using random hexamers and SuperscriptII RT (GibcoBRL). For each allele four overlapping RT-PCR fragments were sequenced to identify mutant lesions. For *t3*, genomic DNA was extracted from individual $cls^{t3}$ mutant embryos, a genomic fragment encoding the N-terminal of the Sox10 protein was amplified by PCR and sequenced commercially (Oswell, Southampton). The *t3* insertion sequence is available in Genbank (Accession Number, AF404490).

### Ectopic expression in zebrafish embryos

The full coding region of $cls^{m618}$ was amplified by RT-PCR using primers *Cla*I-S21 (5′-CCATCGATACCTACCGAAGTCACCTGT-GG-3′) and S27-*Xba*I (5′-GCTCTAGAGTTTGTGTCGATTGTGG-TGC-3′). The 1615bp fragment was subcloned into the *Cla*I/*Xba*I site of the heatshock vector pCSHSP (Halloran et al., 2000) to generate hs>*sox10(L142Q)*. The wild type *sox10* construct, hs>*sox10*, was generated by site-directed mutagenesis of hs>*sox10(L142Q)* using QuikChange Site Directed Mutagenesis Kit (Stratagene). Sequencing confirmed the successful generation of both clones. DNA purified for injection using Microcon Filter Devices (Millipore) was diluted to a concentration of 25 ng/μl, with 0.1% Phenol Red. $cls^{m618}$ or their wild-type siblings were injected with 2 nl of either hs>*sox10* or hs>*sox10(L142Q)* at the one- to two-cell stage and incubated at 28.5°C. As appropriate, embryos were heatshocked twice (at 10-13 hpf and 22-24 hpf), by incubation at 37°C for 1 hour. *cls* embryos were scored for rescue at 48 hpf using a MZ12 dissecting microscope (Leica). Rescue was defined as the presence of at least one melanophore of wild-type morphology; these are never seen in uninjected mutant embryos.

### Iontophoretic labelling and clonal analysis of single neural crest cells

Iontophoretic labelling of individual neural crest cells was performed using essentially the method of Raible et al. (Raible et al., 1992), except that micropipette tips were filled with 3% lysinated rhodamine dextran and 3% biotinylated dextran (both $10\times10^3$ $M_r$; Molecular Probes) mixture dissolved in 0.2 M KCl and the needles backfilled with 0.2 M KCl. Embryos from heterozygotes for *cls* alleles *t3*, *tw2* and *tw11* were used; no phenotypic differences between these alleles were seen, so we do not distinguish them here. Individual NCCs were labelled by intracellular injection with dye. Premigratory cranial crest cells, between the posterior eye and the anterior boundary of somite 1, were labelled in 4-14 somite (11-16 hours post fertilisation, hpf) embryos, while premigratory trunk NCCs at the level of somite 7 were labelled in 16-22 somite (18-20 hpf) embryos. Premigratory trunk NCCs at the level of somite 14 were labelled in 22-25 somite (20-22 hpf) embryos. Embryos were recovered and raised for several hours at 28.5°C in embryo medium with 1% v/v penicillin/streptomycin solution (Gibco). Embryos were remounted and examined to identify those with only a single labelled NCC; only these embryos were analysed.

Labelled cells were monitored using Nomarski and fluorescence optics on a BX50WI microscope (Olympus), and documented on a Eclipse E800 microscope (Nikon) using low light level, video-enhanced fluorescence microscopy. Labelled cells were monitored twice daily for up to 3 days until progeny could be identified using published morphological criteria (Raible and Eisen, 1994; Raible et al., 1992; Schilling and Kimmel, 1994). Thus, melanophores contained melanin granules; xanthophores were visibly yellow and autofluoresced at a wavelength near that of fluorescein; iridophores contained iridescent granules; dorsal root ganglial neurones were identified by position ventrolateral to the neural tube combined with a visible neurite; sympathetic neurones were positioned ventral and lateral to the notochord and showed neurites; Schwann cells were elongated and positioned along axonal processes, e.g. trigeminal nerve; satellite glia were associated with the ganglionic sheath or appeared to be wrapped around neuronal stomata; craniofacial cartilage formed characteristic stacks of vacuolated cells in the jaw or gill arches; fin mesenchyme occupied a position within the dorsal fin fold and showed characteristic asymmetric organisation of projections (Smith et al., 1994). Cells which could not be assigned to any of the described groups, owing to their position in regions of limited optical resolution or lack of distinctive morphologies, were classified as 'unidentified'. In cranial regions, these cells undoubtedly included cells of connective tissue fates (Schilling and Kimmel, 1994).

### TUNEL

TUNEL (terminal deoxynucleotidyl transferase (TdT)-mediated deoxyuridinetriphosphate (dUTP) nick end-labelling) of double-strand DNA fragmentation was used to confirm apoptosis in cells with apoptotic morphology. $cls^{t3}$ and their wild-type siblings were fixed overnight at 4°C in 4% paraformaldehyde and TUNEL performed using fluorescein dUTP and developed using 4-Nitroblue tetrazolium chloride and 5-Bromo-4-chloro-3-indolyl-phosphate (Boehringer Mannheim; Reyes, 1999).

### Morpholino injections

AB wild-type embryos at 25, 30 and 35 hpf were injected with either a high (16.5 ng) or low (9 ng) dose of a morpholino designed to knock-down *sox10*, as described previously (Dutton et al., 2001). Effects on *nac/mitf* expression in melanoblasts at 25 hpf were evaluated by counting *nac/mitf*-expressing cells in one half of the trunk in each of 20 embryos at each dose.

### Whole-mount in situ hybridisation and antibody staining

RNA in situ hybridisation was performed as described by Kelsh and Eisen (Kelsh and Eisen, 2000), on $cls^{m618}$, $cls^{tw11}$ and $cls^{t3}$, their wild-type siblings and morphants. Probes for the following genes were used: *nac/mitf* (Lister et al., 1999), *spa/kit* (Parichy et al., 1999); *dopachrome tautomerase(dct)* (Kelsh et al., 2000b); *dlx2* (Akimenko et al., 1994); *forkhead 6* (*fkd6*) (Odenthal and Nuesslein-Volhard, 1998).

Antibody staining with anti-Hu, mAb 16A11 (Marusich et al., 1994), was performed using peroxidase-antiperoxidase (VECTASTAIN® Elite ABC kit) and DAB substrate.

## RESULTS

### *cls* and a zebrafish *sox10* homologue map to the same region of Linkage Group 3

The strong phenotypic similarity between $sox10^{Dom}$ mice and *cls* mutants suggested a zebrafish *sox10* homologue as a candidate gene for *cls*. We used RT-PCR to clone a partial *sox10*-like HMG box and RACE RT-PCR to clone 5′ and 3′ regions. Sequencing these clones revealed an open reading frame encoding a *sox10* homologue, which we refer to as *sox10* (Fig. 1A,B).

Genetic mapping of 274 meioses placed *cls* on LG 3 within a 3.9 cM interval between markers z872 and z13387 (Fig. 1C). Two oligonucleotide primers amplified a 931 bp fragment from zebrafish, but not from control mouse genomic DNA. Amplification from the DNAs in the LN54 radiation hybrid mapping panel (Hukriede et al., 1999) with these primers mapped *sox10* to LG 3, 0cR from the marker z8492 (LOD score=17.6; Fig. 1C). The striking linkage of *cls* and *sox10*, together with in situ hybridisation experiments showing *sox10* expression in neural crest (see below), strongly supported our hypothesis that *cls* might encode *sox10*.

### *sox10* is disrupted in *cls* mutants

We used RT-PCR to amplify the *sox10*-coding region from 27 hpf homozygous mutants of 3 *cls* alleles. Sequencing these PCR products identified sequence differences from wild-type consistent with them causing the mutant phenotype (Fig. 1D,E). Two alleles show an A→T transversion, resulting in a premature Stop codon; the third is a non-conservative substitution (L142Q) of a fully-conserved residue in the HMG domain. A fourth allele, *t3*, showed highly reduced RNA expression using a 3′ probe for whole-mount RNA in situ hybridisation. PCR from genomic DNA identified a 1.5 kb insertion in all *t3* mutants that was not present in wild types. Sequencing genomic DNA from *t3* homozygotes identified a 1397 bp insertion with sequence homology to a transposon first identified in a mutant *no tail* allele (data not shown; Schulte-Merker et al., 1994). The insertion interrupts the *sox10*-coding sequence upstream of the HMG domain and adds eight novel amino acids before prematurely truncating the protein (Fig. 1D).

To test further whether *cls* encodes *sox10*, we attempted to rescue the *cls* phenotype with ectopic *sox10* expression under heat shock control. We took advantage of the consistent absence of large, stellate melanophores in every *cls* embryo. Ectopic expression of wild-type *sox10* rescued 1-40 melanophores to a wild-type morphology in 48% of $cls^{m618}$ mutants, while ectopic mutant *sox10(L142Q)* failed to do so (Table 1; Fig. 2). Furthermore, while *cls* melanophores always remain dorsal to the neural tube, rescued melanophores frequently migrated, even to very distal positions (Fig. 2).

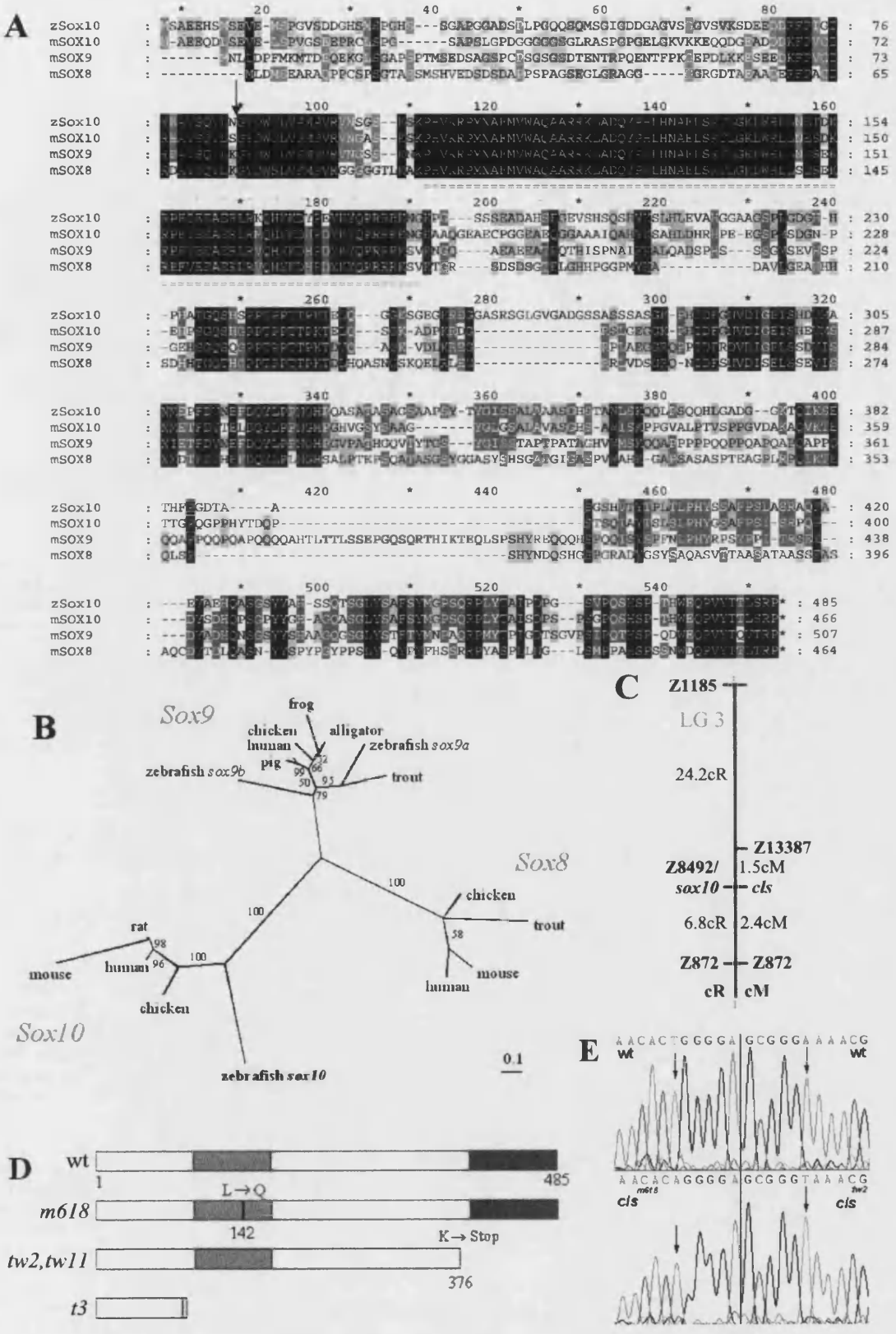

A
zSox10
mSOX10
mSOX9
mSOX8
B
Sox9
frog
chicken
human
alligator
zebrafish sox9a
pig
zebrafish sox9b
trout
Sox8
chicken
trout
rat
human
mouse
mouse
human
chicken
Sox10
zebrafish sox10
0.1
C
Z1185
LG 3
24.2cR
Z13387
Z8492/
sox10
1.5cM
cls
6.8cR
2.4cM
Z872
Z872
cR
cM
D
wt
1
485
m618
L→Q
142
K→Stop
tw2,tw11
376
t3
E
wt
wt
cls
cls

**Fig. 1.** A zebrafish *sox10* homologue maps to the region of the *cls* locus. (A) Sequence comparison of predicted zebrafish Sox10 homologue (44, 45 and 60% identity to mouse Sox8, Sox9 and Sox10, respectively). Blocks of identity corresponding to all proposed functional domains can be seen, including the HMG domain (red underline; 95% amino acid identity), N-terminal synergy domain (1-105; 48% identity), dimerisation domain (66-105; 78% identity), C-terminal transcriptional activation domain (395-485; 76% identity) and a domain C-terminal to the HMG domain corresponding to a putative protein-protein interaction domain (234-325; 64% identity; Bondurand et al., 2000; Kuhlbrodt et al., 1998a; Kuhlbrodt et al., 1998b; Liu et al., 1999; Peirano and Wegner, 2000). (B) Maximum likelihood phylogenetic tree of subgroup E Sox genes. Zebrafish *sox10* clusters within the *Sox10* clade of vertebrate Sox genes. The Accession Numbers for the sequences are as follows: chicken *Sox8* (AF228664); trout *SoxP1* (D83256); mouse *Sox8* (AF191325); human *SOX8* (AF226675); frog *Sox9a* (AB035887); alligator *Sox9* (AF106572); trout *Sox9* (AB006448); zebrafish *sox9a* (AF277096); zebrafish *sox9b* (AF277097); chicken *Sox9* (AB012236); pig *Sox9* (AF029696); human *SOX9* (Z46629); zebrafish *sox10* (AF402677); chicken *Sox10* (AF152356); mouse *Sox10* (AF047389); rat *Sox10* (AJ001029); human *SOX10* (NM_006941). (C) Mapping using the LN54 panel placed *sox10* on LG 3 in the region of the *cls* locus identified using microsatellite markers (we found four recombinants between *cls* and z13387 in 274 meioses). Note that z8492 was not polymorphic and could not be analysed in the mapping cross. (D) Schematic to illustrate changes in Sox10 mutant proteins. In $cls^{m618}$ a T425A substitution results in a non-conservative change (Leu142Gln) within the HMG domain (red). In $cls^{tw2}$ and $cls^{tw11}$, a A1126T substitution introduced a Stop codon truncating the protein just N-terminal to the transactivation domain (blue). Insertion of a 1.4 kb transposon at the site indicated by the arrow in A disrupts *sox10* in $cls^{t3}$ and introduces a C-terminal extension of eight novel amino acids before premature truncation N-terminal to the HMG domain (yellow). (E) Chromatogram traces to show nucleotide changes affecting *sox10*-coding regions in $cls^{m618}$ and $cls^{tw2}$.

**Table 1. Rescue of *cls* phenotype using ectopic *sox10* expression**

| Construct | Heat shock | Injected $cls^-$ that survived | Rescued $cls^-$ (% rescued) | Mean number of melanophores per rescued embryo |
|---|---|---|---|---|
| *hs>sox10* | + | 92 | 44 (48) | 15 |
| *hs>sox10* | – | 48 | 17 (35)* | 4.3 |
| *hs>sox10(L142Q)* | + | 122 | 0 (0) | 0 |

*A similar degree of leakiness with this promoter has been reported by Lister et al. (Lister et al., 1999); note that the degree of rescue is much less in the absence of heat shock.

### *sox10* expression is disrupted in *cls* mutants

We used in situ hybridisation to examine *sox10* expression in wild-type and *cls* embryos. In wild types, expression was first detected at the one-somite stage in cells in the lateral neural plate (data not shown). Throughout somitogenesis stages, strong *sox10* expression was seen in premigratory NCCs and extended progressively more caudally in older stages (Fig. 3A-C,F). Double RNA in situ hybridisation studies show that there is extensive, but incomplete, overlap of *sox10* and *fkd6*, a marker expressed widely in premigratory NCCs (Fig. 3O,P) (Odenthal and Nuesslein-Volhard, 1998). *sox10* expression was maintained in some migrating NCCs on the medial migration pathway (Fig. 3C,F). By 30-40 hpf *sox10* expression on the medial pathway is organised in segmentally arranged clusters adjacent to the notochord, presumably developing Schwann cells associated with the segmental nerves (Fig. 3Q); *sox10* expression is lost from this site by 48 hpf. Although NCCs are found extensively on the lateral pathway from 24 hpf (Raible et al., 1992), counts of *sox10*-expressing cells show that expression was essentially absent from cells on this pathway (Fig. 3M,N). This demonstrates the rapid downregulation of *sox10* from pigment cell precursors, as NCCs on this pathway form only pigment cells (Raible and Eisen, 1994). Consistent with this, xanthophores and most pigmented melanophores show no detectable *sox10* expression (Fig. 3L), although some weakly expressing melanophores were noted at earlier stages (Fig. 3K). Expression was not seen in fin mesenchyme nor in *dlx2*-expressing craniofacial cartilage precursors (Fig. 3D,E), although differentiating jaw cartilage shows weak expression by 60 hpf (data not shown). Cells expressing *sox10* accumulated in clusters corresponding to the forming cranial ganglia and extending along the posterior lateral line nerve by 24 hpf (Fig. 3F,H), in a pattern reminiscent of *fkd6* expression (Kelsh et al., 2000a), suggesting that *sox10* is expressed in satellite glial and Schwann cells. Consistent with this interpretation, double labelling with anti-Hu antibody confirmed the non-overlapping expression of *sox10* and this neuronal marker (Fig. 3J). *sox10* is maintained in developing Schwann cells on the posterior lateral line nerve up to 60 hpf (data not shown). *sox10* expression was prominent in enteric nervous system precursors at 60 hpf (Fig. 3S).

At early stages, premigratory crest showed equivalent patterns of *sox10*-positive cells in wild-type and $cls^{tw11}$ and $cls^{m618}$ mutant embryos. Counts at 24, 30 and 35 hpf revealed that although in wild type *sox10*-expressing cells were rapidly lost from the premigratory position, they remained here in $cls^{m618}$ mutants; by contrast, *sox10* expression in migrating trunk NCC was broadly comparable between wild type and $cls^{m618}$ mutants (Fig. 3M,N). Additionally, *cls* mutants were

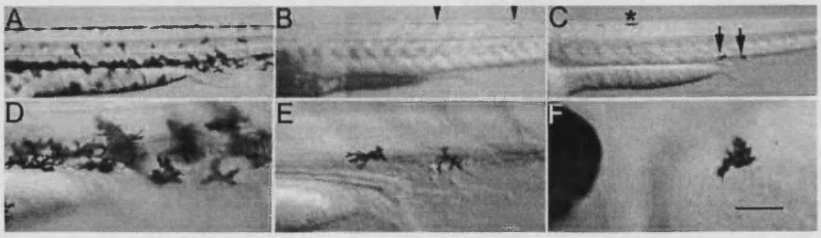



**Fig. 2.** *cls* phenotype is rescued by ectopic *sox10* expression. Wild-type embryos show many large, strongly pigmented melanophores at 48 hpf (A; close-up of anal region in D), while hs>*sox10(L142Q)*-injected $cls^{m618}$ mutants (B) show only tiny melanised spots in position of premigratory NCCs (arrowheads). (C,E,F) By contrast, $cls^{m618}$ embryos injected with hs>*sox10* and heat-shocked, show mosaic rescue of melanophores. This embryo showed one rescued melanophore in the dorsal stripe (* in C), two in the ventral stripe (arrows in C; close-up in E) and one on the yolk sac (F). Scale bar: 125 μm in A-C; 70 μm in D-F.

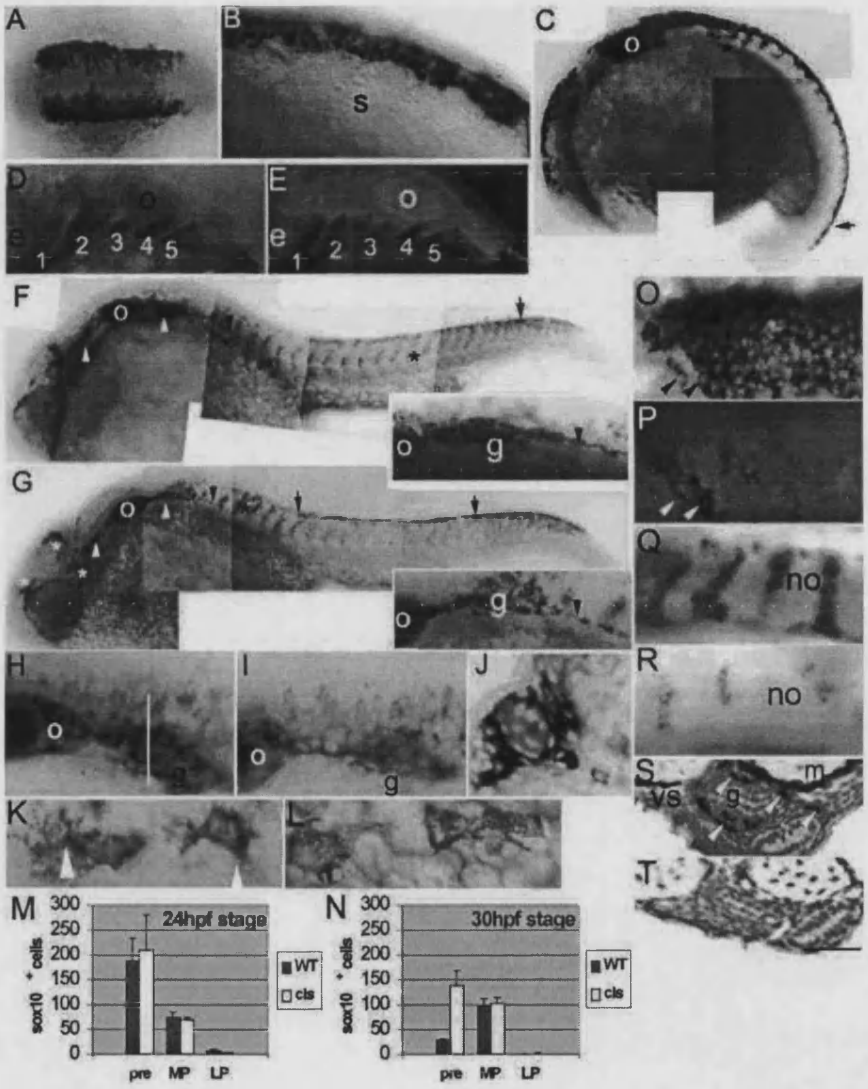



**Fig. 3.** Embryonic *sox10* expression in wild-type and *cls* embryos. (A,B) *sox10* is expressed in most cranial (A; five-somite stage) and trunk (B; 14-somite stage) premigratory NCCs; double mRNA in situ hybridisation with *fkd6* (O, purple) and *sox10* (P, green) in a six-somite stage embryo reveals extensive overlap, although some individual NCCs lack *sox10* (arrowheads). (C) 18-somite stage embryo shows strong expression in premigratory (black arrow) and migrating (asterisk) NCCs and otic vesicle (o). (D,E) Double mRNA in situ hybridisation with *dlx2* (D, purple) and *sox10* (E, green) in 29 hpf stage embryos shows absence of *sox10* expression in developing branchial arches (1-5). (F) By 24 hpf in wild types, strong expression is associated with cranial ganglia (white arrowheads) and posterior lateral line nerve (black arrowhead; enlarged in inset), otic vesicle (o), migrating NCCs throughout trunk (asterisks) and in premigratory crest (arrow). (G) At 24 hpf in *cls* mutants, expression in the head is clustered (white asterisks) and cranial ganglia have reduced labelling (white arrowheads). Cells expressing *sox10* extend along the posterior lateral line nerve (arrowhead and inset). Rostral trunk shows some migrating cells (black asterisk), but *sox10*-positive cells are clustered dorsal to the neural tube (arrows) in trunk and tail. (H,I) Combined *sox10* in situ hybridisation (purple) and anti-Hu antibody labelling (orange) shows strong *sox10* expression associated with wild-type (H) posterior lateral line ganglion (g), much reduced in *cls* mutant (I). (J) In transverse section of wild-type ganglion (approximate position indicated by white line in H), *sox10* expression is strong peripherally (non-neuronal cells), but absent centrally (neurones). (K,L) 36 hpf wild-type embryos show weak expression in some melanophores, but not all. Thus, weak expression (arrowhead) is seen in some cells of the dorsal stripe (K), but not in cells on the yolk sac (L). (M,N) Number of *sox10*-expressing NCCs in different locations (premigratory, pre; migrating on medial pathway, MP; migrating on lateral pathway, LP) of trunk and anterior tail (somites 1-20) at 24 (M) and 30 hpf (N) in WT and *cls* mutants. (Q,R) Segmentally arranged lines of *sox10*-positive cells lying adjacent to the notochord (no), presumably glia, are abundant in wild type (Q) and only weakly affected in *cls* mutants (R) at 40 hpf. (S,T) Transverse section of trunk of 60 hpf wild-type embryo (S) shows enteric nervous system expression (arrowheads) lateral to the gut (g), absent in *cls* mutant (T). e, eye; m, muscle; s, somite; vs, ventral stripe melanophores. All images are lateral views, rostral towards the left, dorsal upwards, except dorsal views of A,K,O,P. Scale bar: 20 μm in A; 50 μm in B,H-I,K-L; 120 μm in C; 75 μm in D,E; 150 μm in F,G; 35 μm in J; 65 μm in O,P; 55 μm in Q,R; 45 μm in S,T.

distinguishable from 24 hpf by the clustered, not scattered, distribution of *sox10*-positive cells in the head (Fig. 3F,G) and a variable reduction in the number of *sox10*-positive cells in cranial ganglia and on cranial nerves (Fig. 3H,I). By 35 hpf, *cls* mutants were readily distinguished from wild types by their reduced number of *sox10*-positive cells, concentrated in a premigratory position dorsal to the neural tube or clustered near the posterior lateral line ganglia; putative Schwann cells were missing from the posterior lateral line nerve. *cls* mutants showed reduced expression in the putative Schwann cells found as segmental clusters of *sox10*-positive cells (Fig. 3Q,R). Enteric nervous system precursors were absent at 60 hpf (Fig. 3T) in *cls* mutants.

In contrast to the other mutant alleles examined, $cls^{t3}$ mutants consistently show highly reduced *sox10* transcripts when examined using the 3′ *sox10* probe at all stages. When examined using a probe lying 5′ to the insertion site, *sox10* expression levels were comparable with those in other mutant

**Table 2. Fates of single neural crest cells injected with lineage tracer**

| | Trunk (somite 14) | | | | Cranial* | | | |
|---|---|---|---|---|---|---|---|---|
| | Wild type | | *colourless* | | Wild type | | *colourless* | |
| Fate‡ | Number | Percentage | Number | Percentage | Number | Percentage | Number | Percentage |
| Ectomesenchymal | | | | | | | | |
| Cartilage | – | – | – | – | 6 | 7 | 2 | 6 |
| Fin mesenchyme | 5 | 7 | 2 | 8 | – | – | – | – |
| Non-ectomesenchymal | | | | | | | | |
| *Pigment* | | | | | | | | |
| Melanophore | 26 | 34 | – | – | 12 | 14 | – | – |
| Xanthophore | 9 | 12 | 4§ | 16 | 12 | 14 | 7§ | 21 |
| Iridophore | 3 | 4 | – | – | 2 | 2 | – | – |
| Mixed pigment | 10 | 13 | – | – | – | – | – | – |
| *Neural* | | | | | | | | |
| Cranial ganglia | – | – | – | – | 9 | 11 | – | – |
| Dorsal root ganglia | 6 | 8 | – | – | – | – | – | – |
| Sympathetic neuron | 3 | 4 | – | – | – | – | – | – |
| Schwann cell | 1 | 1 | – | – | – | – | – | – |
| Mixed neural¶ | 5 | 7 | – | – | – | – | – | – |
| *Mixed pigment/neural* | | | | | | | | |
| Melanophore + Dorsal root ganglia | 1 | 1 | – | – | – | – | – | – |
| Other | | | | | | | | |
| Died | 1 | 1 | 19§ | 76 | 11 | 13 | 15§ | 45 |
| Unidentified | 5 | 7 | – | – | 33 | 39 | 9 | 27 |
| Total | 75 | 100 | 25 | 100 | 85 | 100 | 33 | 100 |

*All progeny within a single clone adopted the same fate in every case, as shown previously (Schilling and Kimmel, 1994).

‡Only single injected cells that survived and were identified as NCCs in the afternoon on the day of labelling are included. The numbers represent the number of clones whose cells adopted the indicated fates, using the criteria outlined in the Materials and Methods.

§All xanthophores identified in *cls* embryos subsequently died; hence in total, 23 (92%) and 22 (66%) cells died in the somite 14 and cranial samples.

¶Mixed neural fates include clones with both neuronal and glial derivatives.

alleles, suggesting that insertion of the transposon does not disrupt transcription of 5′ sequences (data not shown).

Outside neural crest, *sox10* expression was particularly strong in otic placode and otic vesicle, and from 11-somite stage onwards (Fig. 3C,F,H), was detected in pectoral fin and in some spinal cord cells from 36 hpf. Expression in the ear was significantly weaker in *cls* mutants by 40 hpf (data not shown).

### *cls* non-ectomesenchymal neural crest cells die prior to differentiation

Previous characterisation of *cls* embryos catalogued a strong reduction in non-ectomesenchymal neural crest derivatives (Kelsh and Eisen, 2000). To analyse the cell biological basis for loss of these neural crest derivatives, we used iontophoretic labelling of single NCCs. We labelled premigratory NCCs in *cls* mutants and their wild-type siblings in two regions that generate ectomesenchymal fates, and scored them for the fate(s) adopted by their progeny (Table 2). In wild-type embryos, almost all (148/160; 93%) labelled cells survived throughout the experiment, and all major derivatives, both ectomesenchymal and non-ectomesenchymal, were identified among the clones. Consistent with our previous analyses, labelled cells generated a similar proportion of ectomesenchymal fates in *cls* mutants (4/58; 7%) and wild-type siblings (11/160; 7%). Furthermore, even in *cls* mutants, cells in these clones migrated and differentiated normally. In wild-type embryos, most cells where fates were identifiable (111/122; 91%) differentiated into recognisable non-ectomesenchymal derivatives. By contrast, most identifiable clones in *cls* embryos (45/49; 92%) died late on the second day (~35-45 hpf). Most showed no sign of morphological differentiation or pigmentation (34/49; 69%). The only exception, clones that developed some xanthophore pigmentation (11/49; 22%), always had abnormal morphology, being rather rounded and blebbed, and died soon after differentiation (Fig. 4N). A series of single cell injections in NCCs at the axial level of somite 7 gave similar results, with 92% of *cls* cells dying by around 48 hpf, in contrast to only 2% in wild-type siblings (data not shown); note that at this axial position no ectomesenchymal fates have been reported (Raible and Eisen, 1994). Our results are consistent with the severe reduction in differentiated non-ectomesenchymal fates in *cls* embryos and suggest that extensive NCC death is associated with the *cls* phenotype.

### Non-ectomesenchymal crest derivatives die by apoptosis during a discrete time-window

The morphological appearance of dying cells in our clonal analyses suggested an apoptotic mechanism of death. We combined TUNEL with biotin-detection of the labelled clone in each of five *cls* embryos and showed that the dying cells had fragmented DNA (Fig. 4A-C). We conclude that many non-ectomesenchymal clones in *cls* embryos die by an apoptotic mechanism.

Analysis of cell survival in our cranial NCC data set shows that death of labelled NCCs occurs within a relatively discrete time-window (Fig. 4L). In both *cls* and wild-type embryos, around 10% of labelled clones had died by the morning after injection. In wild-type embryos, the number of surviving

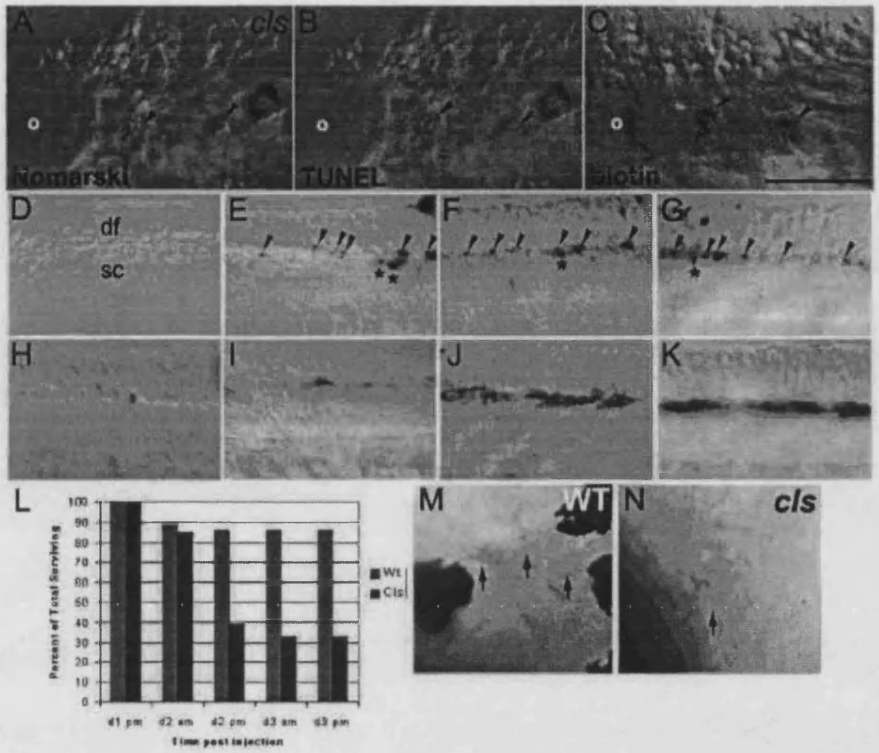



**Fig. 4.** NCC death in *cls* embryos. (A-C) NCC clones die by an apoptotic mechanism in *cls* embryos. Lateral views of 40 hpf *cls* embryo in which two daughters of a single labelled NCC contributed to the posterior lateral line ganglion, lying just posterior to the otic vesicle (o). In the live embryo, both cells show blebbed morphology typical of apoptotic cells when viewed with Nomarski optics (A). After fixation and processing for TUNEL (B) and detection of the biotinylated-dextran lineage-tracer (C), visible TUNEL signal of these clonal cells indicates DNA fragmentation characteristic of apoptotic cells. (D-K) Whole-mount TUNEL shows NCC apoptosis in *cls* embryos. Lateral views of dorsal spinal cord (sc) in tail of 30 (D,H), 35 (E,I), 40 (F,J) and 45 (G,K) hpf embryos show apoptotic NCCs immediately dorsal to the spinal cord from 35 hpf in *cls* (arrowheads, D-G), but not wild-type (H-K), embryos. Scattered TUNEL-positive cells are prominent in dorsal spinal cord (*) of *cls* embryos (E-G); these are occasionally seen in wild-type siblings at these stages (data not shown). df, dorsal fin. (L) Time-course of labelled single cranial NCC clone survival in *cls* mutants and their wild-type siblings. Percentage of surviving clones is given at each of the five standard time points when embryos were examined. The time points correspond to approximately 16, 32, 40, 56 and 64 hpf, respectively. The first time point includes only single labelled NCCs based on examination within a few hours after labelling. See text for further details. (M,N) Wild-type xanthophores (arrows) at 48 hpf have a very flattened, thin morphology and are only weakly colored (M), while a dying *cls* xanthophore (arrow) shows characteristic apoptotic morphology and concentrated yellow coloration, which was usually visible by 35 hpf (N). Scale bar: 100 μm in A-C; 50 μm in D-K; 75 μm in M,N.

clones was essentially unchanged at later times. However, in *cls* mutants, the number of surviving clones dropped precipitously to only 40% within the period of around 5-10 hours between the two observational time-points on the second day after labelling (equivalent to ~40 hpf), but then remained constant. We interpret the early (overnight after injection) loss of wild-type and *cls* clones as representing death due to damage during labelling. The subsequent late loss of clones in *cls* embryos corresponds to the time of appearance of apoptotic cells and we attribute this to the *cls* phenotype.

To examine the extent to which apoptosis contributes to the *cls* phenotype, we used whole-mount TUNEL to examine apoptosis of NCCs (Fig. 4D-K). We saw a notable concentration of apoptotic cells dorsal or dorsolateral to the neural tube in *cls* embryos, but not in wild-type control embryos. A timecourse of TUNEL between 20 and 60 hpf indicated that cell death in NCCs becomes apparent by 35 hpf and continues to approximately 45 hpf.

### *cls* neural crest cells fail to migrate before undergoing apoptosis

Our in vivo clonal studies revealed that, while wild-type NCCs migrated extensively on both medial and lateral pathways, most *cls* NCCs failed to leave the premigratory crest area. Thus, excluding ectomesenchymal derivatives (which always migrated normally), only 2/31 (6%) of cranial NCCs appeared to migrate away from their initial positions. At the level of somite 14, all wild-type labelled cells left the premigratory area, and more than 75% migrated at least 50 μm from their original position. Only fin mesenchyme clones migrated normally in *cls* mutants. Of the remaining clones, only 4/23 (17%) migrated away from the premigratory crest; another cell extended towards the horizontal myoseptum, but maintained contact with the premigratory area. Of these five cells, two migrated on the lateral pathway, but underwent apoptosis after moving about half way to the horizontal myoseptum. The other three cells migrated on the medial path. In one case, the cell did not divide, migrated to a position appropriate for a dorsal root ganglion, then died before overt differentiation. The other two cells divided once; in each clone, one sister cell remained dorsal to the neural tube, and may have been an extramedullary cell (Kelsh and Eisen, 2000), while the other migrated to a dorsal root ganglial position, but failed to undergo axonogenesis, consistent with the impaired touch response of 5 dpf larvae. We obtained similar results for labelled NCCs from the region of somite 7 (data not shown). This direct observation of failed NCC migration in *cls* embryos is consistent with our observation of *dct*-expressing melanoblasts and *sox10*-expressing NCCs concentrated dorsal to the neural tube in mutant embryos (Kelsh et al., 2000b) (this work).

### *cls* mutants and *sox10* morphants fail to express genes critical for melanophore specification and migration

*cls* mutants combine an extensive failure of NCC migration with late apoptotic death of non-ectomesenchymal derivatives. Three recent studies suggested a possible molecular genetic mechanism for these aspects of the phenotype. *sparse* (*spa*), a *kit* homologue, is crucial for survival and migration of melanophores and melanoblasts (Parichy et al., 1999; Kelsh et al., 2000b) and *nacre* (*nac*), a *microphthalmia transcription factor* (*mitf*) homologue, is crucial for melanophore specification and required for *spa* expression (Lister et al., 1999). We used in situ hybridisation at 20-35 hpf to ask whether *spa* or *nac* expression was disrupted in *cls* embryos (Fig. 5). We found that both *spa* and *nac* expression were absent from NCCs, even at a time when many *dct*-positive cells are present, suggesting that the *cls* melanophore phenotype might result from loss of *nac* expression. By contrast, *spa* expression in intermediate cell mass and *nac* expression in the pigmented retinal epithelium are unaffected in *cls* mutants.

We have recently shown that injection of *sox10* morpholino oligonucleotides phenocopies the *cls* mutant phenotype (Dutton et al., 2001). The morphant phenotype varies depending upon the amount of morpholino injected, with high doses phenocopying all aspects of the phenotype of the strong *cls* alleles. Injection of lower doses results in a phenocopy of the weak *cls* allele phenotype, with embryos showing some melanophores, concentrated in dorsal positions (Dutton et al., 2001; Malicki et al., 1996). The number of *nac/mitf*-expressing cells in embryos injected with a low dose of *sox10* morpholino (median number=78; $n$=20) is significantly higher than in embryos injected with a high dose (median number=20; $n$=20; Mann-Whitney U-test, $P<0.05$; Fig. 5C,D), consistent with the differences in melanophore phenotype.

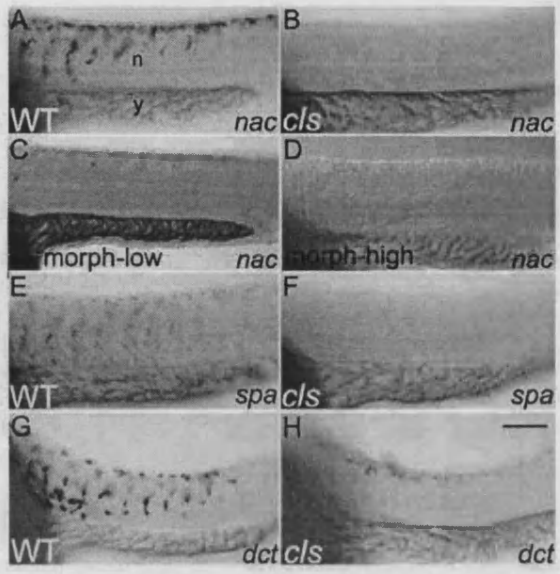



**Fig. 5.** *cls* mutants lack *nac* and *spa* expression. Lateral views of caudal trunk of 25 hpf wild-type and *cls* embryos after in situ hybridisation with *nac* (A,B), *spa* (E,F) and *dct* (G,H) probes. (C,D) *nac/mitf* expression is decreased weakly (C) or strongly (D) after injection with either a low or high dose, respectively, of *sox10* morpholino. n, neural tube; y, yolk sac. Scale bar: 75μm.

## DISCUSSION

Studies in mice have led to identification of *Sox10* as a key gene in human Waardenburg-Shah syndrome. We have shown that the zebrafish *cls* locus is a *sox10* homologue predicted to encode a protein with all the major domains identified in mammals (Fig. 1). We have identified the molecular lesion likely to cause the mutant phenotype in 4 *cls* alleles. In two cases, *tw2* and *tw11*, both originating from the same mutagenised male (Haffter et al., 1996; Kelsh et al., 1996), the lesions are identical and are presumably independent isolations of the same allele. The zebrafish Lys376Stop lesions resemble the human *059* allele (delGA) in lacking the transactivation domain, although they lack the C-terminal extension present in the human allele. In transient transfection assays human 059 mutant protein has been shown to have no transcriptional activation (Bondurand et al., 2000; Kuhlbrodt et al., 1998b; Liu et al., 1999). The *m618* allele substitutes Leu142 in the second alpha helix of the HMG domain and there is no similar mammalian mutation. The *t3* allele results in a severely truncated protein that lacks both the DNA-binding and transcriptional activation domains. It is thus reminiscent of the human Y83X allele which has been proposed to be a functional null (Kuhlbrodt et al., 1998b; Potterf et al., 2000). All four mutant alleles described here show similar, strong phenotypes (Kelsh et al., 1996; Malicki et al., 1996) (data not shown). These considerations, combined with the similarity of these phenotypes with the maximal morphant phenocopy generated with a *sox10* morpholino, lead us to suggest that these alleles are all likely null alleles. Analysis of the activities of these mutant proteins will be interesting to test this proposal.

Zebrafish *sox10* expression is consistent with cell types affected in *cls* mutants and strongly reminiscent of that in mammals. We have taken advantage of the higher resolution of such studies in zebrafish, to define more precisely the extent of *sox10* expression throughout the neural crest at different stages. In wild type, expression is extensive in premigratory neural crest, but comparison with *fkd6* expression shows that a minority of premigratory NCCs lack *sox10* expression. Expression persists in some migrating cells on the medial migration pathway, but is rapidly downregulated in differentiating pigment cells. Glia of the developing peripheral nervous system show strong expression, as has also been described in mice (Kuhlbrodt et al., 1998a; Britsch et al., 2001). Crest cells in forming branchial arches and fin mesenchyme do not show expression, in agreement with mouse, but not human, expression studies (Bondurand et al., 1998; Kuhlbrodt et al., 1998a; Southard-Smith et al., 1998). Late expression in cranial cartilages is intriguing, but will require further study to evaluate its role. In *cls* mutants, *sox10*-expression in premigratory NCCs is initially unaffected but, unlike in wild-type siblings, *sox10*-expressing cells soon accumulate in this position. In the trunk *sox10*-expressing cells migrating on the medial pathway are seen and presumably contribute to the neurones and glia of the dorsal root ganglia, which are detectable in mutants (Kelsh and Eisen, 2000) (this

study). NCCs tend to become clustered in premigratory positions in *cls* mutants, consistent with the observed defect in NCC migration revealed by our single-cell labelling studies.

Outside the neural crest, *sox10* expression is also largely conserved. Thus, expression in the developing inner ear epithelium is seen in zebrafish as well as in mammals (Bondurand et al., 1998; Southard-Smith et al., 1998). In zebrafish, this expression is remarkably strong and persistent in wild type, but is downregulated in *cls* mutants by 40 hpf. As pronounced otic defects are morphologically detectable in *cls* mutants by 48 hpf (R. N. K., unpublished), we suggest that *sox10* is crucial for development of otic epithelium. Limited expression in embryonic central nervous system is seen in zebrafish and mice (Kuhlbrodt et al., 1998a), although detailed studies will be required to establish the cell types involved in each case.

Our single cell labelling studies at three rostrocaudal positions make clear that defects in *cls* mutants are not limited to one axial position, but instead affect non-ectomesenchymal rather than ectomesenchymal crest derivatives. Most NCCs in *cls* mutants show restricted migration and poor or no differentiation before dying by an apoptotic mechanism within a discrete time-window. We interpret the dying cells as being those that would in wild-type siblings form non-ectomesenchymal fates. Thus, in *cls* embryos, NCCs that would normally yield these missing neural crest derivatives are present in premigratory neural crest in normal numbers, but then die before differentiating, usually without migrating. This confirms and extends reports that putative NCCs apoptose during migration in mouse *sox10* mutants (Kapur, 1999; Southard-Smith et al., 1998).

The *sox10* expression pattern in *cls* mutants may appear contradictory to our single cell label results, as the expression studies show normal numbers of *sox10*-positive NCCs migrating on the medial pathway, while the single cell labelling studies show a strong migration defect in the neural crest. However, our data suggest that *sox10* is expressed in almost all premigratory cells, but is rapidly downregulated in ectomesenchymal and pigment cell precursors as they start to migrate. Thus, *sox10* expression in migrating cells reflects just the neural precursors. Our studies show that these cells are much less defective in migration in *cls* mutants. Consistent with this, the labelled cells in *cls* mutants that did migrate normally all took the medial migration pathway and, if they survived, adopted a position consistent with a dorsal root ganglial fate. By contrast, most labelled NCC clones generated pigment cells, consistent with former studies in zebrafish. These pigment cells are severely defective in migration, as shown by molecular markers for xanthophore (*gdh*; Parichy et al., 2000) and melanophore (*dct*; Fig. 5) cell fates (M. Hawkins and R. N. K., unpublished) (Kelsh and Eisen, 2000). Further work will be required to investigate the specification status of the migrating neural precursors in wild-type and mutant embryos.

We have demonstrated the complexity of NCC defects in *cls* embryos, with survival, migration and differentiation of non-ectomesenchymal cell types all affected. Further, a small decrease in clone size suggests that proliferation may also be somewhat affected (K. A. D. and R. N. K., unpublished). This spectrum of defects is not readily consistent with a primary function for wild-type *cls/sox10* in NCC survival, proliferation or differentiation. Although a dying cell might show abnormal migration and proliferation, the timing of appearance of defects in *cls* mutants conflicts with cell survival being the primary defect. We can first distinguish *cls* mutants at 20 hpf (by lack of *nac/mitf* expression), but do not see NCC apoptosis by TUNEL until around 35 hpf; previous work in zebrafish has suggested that the delay between induction of apoptosis and detectable morphological changes and DNA fragmentation is at most 3-4 hours (Ikegami et al., 1999). Likewise the *cls* mutant phenotype cannot be explained as primarily a failure of migration, with subsequent failure of exposure to required trophic factors, because for three subsets of affected cells (three pigment cell types in the stripe dorsal to the neural tube), the premigratory position is also a final location and hence the necessary trophic factors must be available. Despite this, in *cls* mutants all pigment cells, including those of the dorsal stripe, fail to develop properly. Instead, we propose that the primary role of *cls* gene function is in specification of non-ectomesenchymal cell fates; all other defects in *cls* mutants would then be secondary effects of a failure to become properly specified. Thus, in *cls* mutants NCCs are unable to adopt non-ectomesenchymal fates and so fail to differentiate. This differentiation process would include expression of growth and trophic factor receptors, and hence secondary defects might include impaired proliferation and later apoptosis. Our identification of the *cls* gene as a *sox10* transcription factor gene is consistent with this interpretation.

Indeed, for the specific case of melanophore fate, data presented here and elsewhere permit us to propose a molecular model consistent with this interpretation (Fig. 6). In both zebrafish and mammals, Nac/Mitf homologues have been shown to be important transcription factors for specifying melanophore fate (Lister et al., 1999; Opdecamp et al., 1997; Tachibana et al., 1996). We show here that *cls* mutants lack expression of *nac* (indeed it is the earliest defect we have identified). Thus, *sox10* is required for *nac* expression and thus for melanophore lineage specification. In addition, in *nac* mutants, *nac*-positive cells show a failure to migrate highly reminiscent of the *cls* melanoblast phenotype (Lister et al., 1999). Likewise, in both *cls* and *nac* mutants, *spa/kit* expression is lost in melanoblasts. *spa* function is necessary for both melanoblast migration and melanoblast and melanophore survival (Kelsh et al., 2000b; Parichy et al., 1999). Furthermore, the timing of loss of melanoblasts in *spa* mutants, beginning between 30 and 36 hpf, is consistent with the timing of decrease in melanoblast numbers and with the timing of NCC apoptosis in *cls* mutants (Kelsh et al., 2000b; this study). Hence, a simple model can largely explain the *cls* melanophore phenotype. *cls/sox10* has a key role, by direct or indirect activation of *nac/mitf*, in melanophore specification and consequently *cls* mutants lack expression of genes activated by nac, including *spa/kit*, mediating survival and migration, and melanogenic enzymes, mediating pigmentation. In support of this model, injection of *sox10* mRNA into early embryos results in ectopic *nac/mitf* expression and expression of *nac* in *cls* embryos is sufficient to rescue melanophores (S. E. and R. N. K., unpublished). It will be of interest to expand this model to include other targets of *sox10* and *nac*, as well as to extend this model to other non-ectomesenchymal neural crest derivatives affected in *cls* mutants. Our model predicts that the immediate targets of *sox10* will be either master regulators of

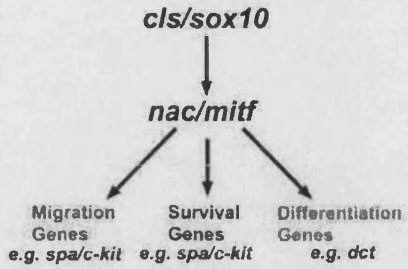



**Fig. 6.** Model of role of *cls/sox10* in melanophore specification. Formal genetic interactions between selected genes known to function in zebrafish melanophore development are schematised. In *cls* mutants, failure to activate *nac* expression (and thus to specify melanophore fate) results in absence of gene products critical for melanophore survival, migration and differentiation. For details, see main text.

neural crest fate (e.g. *nac*) or a multitude of diverse gene products required within specific crest derivatives for the cellular functions of trophic support, migration and differentiation. Thus, it will be important to identify the direct targets of Sox10 transcriptional regulation.

This specific model is likely to be applicable to mammalian neural crest too. Thus, Sox10 directly regulates *Mitf* expression in mouse (Bondurand et al., 2000; Lee et al., 2000; Potterf et al., 2000; Verastegui et al., 2000). Although *Dct*-positive melanoblasts are not detectable in $Sox10^{Dom}$ mice, some evidence of impaired or delayed migration of some NCCs has been presented (Southard-Smith et al., 1998). Similarly, *Kit* function is critically involved in promoting melanoblast migration and survival, is rapidly lost from melanoblasts in *Mitf* mutants and is known to be a direct target of Mitf, at least in cultured mast cells (Opdecamp et al., 1997; Wehrle-Haller and Weston, 1995; Wehrle-Haller and Weston, 1997). In *Sox10* mutant mice, *Kit* expression in melanoblasts may be absent, although the presence of unaffected cells in the skin, suggested to be mast cells, means that this result needs to be confirmed (Bondurand et al., 2000).

*sox10* alone cannot be sufficient to specify each of the non-ectomesenchymal fates. We believe that *sox10*, while crucial for each of them, is not sufficient; instead, we propose that *sox10* is one of a combination of genes that together specify non-ectomesenchymal fates from neural crest. This model helps explain why cells that express molecular markers for affected cell types can be identified in *cls* mutants (Kelsh and Eisen, 2000). For example, although melanophores barely express melanin, xanthophores show much pigment in *cls* mutants. If a combination of transcription factors together specify a cell type, each factor might regulate only partially overlapping subsets of the characteristics of individual fates. In the absence of *sox10* gene function, those aspects of melanophore and xanthophore fate requiring *sox10* function would be affected (e.g. synthesis of melanin and trophic machinery, respectively), while others would remain at least partly unaffected (e.g. *dct* expression and pteridine pigment synthesis, respectively). A directly analogous mechanism has been proposed for autonomic neurone specification from the neural crest, with different genes regulating pan-neuronal and subtype specific cell properties (Anderson et al., 1997). Factors that interact with Sox10 to regulate target genes have been described, including the paired box transcription factor Pax3 and the POU-domain protein Tst-1/Oct6/SCIP (Bondurand et al., 2000; Kuhlbrodt et al., 1998a). However, it remains to be seen whether these co-factors have any role in distinguishing non-ectomesenchymal fates. Identification of combinations of transcription factors required to specify each fate will be a promising line of future research.

Our specification model predicts that *sox10* function is required only in NCCs fated to non-ectomesenchymal derivatives. *fkd6* is expressed very broadly in premigratory NCCs (Odenthal and Nuesslein-Volhard, 1998) and shows extensive, but incomplete, overlap with *sox10*. Further studies will be required to test whether *sox10-; fkd6+* NCCs represent ectomesenchymal precursors. However, our results do show that, consistent with our model, expression is, at least, rapidly downregulated in cells that adopt ectomesenchymal fates. In mammals, *Sox10* has sometimes been assumed to be a generic NCC marker (Pattyn et al., 1999), but our results show that, in fish, overlap with other NC markers is incomplete, even at premigratory stages.

In many respects, zebrafish neural crest development is typical of neural crest development in all vertebrates (Raible et al., 1992). In contrast to mouse and human *Sox10* mutations, none of the fish *sox10* mutations show dominant effects on melanophore number or pattern (R. N. K., unpublished). In mice, expressivity of the *sox10* haploinsufficiency phenotype is dependent on the genetic background (Lane and Liu, 1984). Presumably, in the AB background in zebrafish heterozygotes have sufficient Sox10 for normal melanophore development. Nevertheless, the conserved expression pattern and homozygous phenotypes in the two species suggest that the detailed cell-biological basis for the *cls* phenotype proposed here illuminates a possible mechanism behind Waardenburg-Shah syndrome and Hirschsprung's disease, in which reduced SOX10 function may reduce the number of melanoblasts and enteric precursors specified.

We thank J. James for fish husbandry, L. D. Hurst for generation of Fig. 1B, D. Raible for pCSHSP, and C. Miller, D. Parichy and D. Raible for cDNA clones. We also gratefully acknowledge J. Eisen, D. Raible, J. M. W. Slack, R. Adams, W. Bennett, A. Ward and J. R. Dutton for helpful criticism and discussion of the manuscript. R. G. thanks S. Rudolph-Geiger for technical assistance. This work was supported by a Praxis XXI PhD Studentship (S. S. L.), a University of Bath studentship and ORS award (T. J. C.), a Royal Society Research Grant, Wellcome Trust Project Grant 050602/Z, a MRC Co-operative Group Component Grant G9810985, and MRC Co-operative Group Grant G9721903. R. G. and P. H. were supported by the German Human Genome Project (DHGP Grant 01 KW 9919).

## REFERENCES

**Akimenko, M. A., Ekker, M., Wegner, J., Lin, W. and Westerfield, M.** (1994). Combinatorial expression of three zebrafish genes related to distal-less: part of a homeobox gene code for the head. *J. Neurosci.* **14**, 3475-3486.

**Anderson, D. J., Groves, A., Lo, L., Rao, M., Shah, N. M. and Sommer, L.** (1997). Cell lineage determination and the control of neuronal identity in the crest. *Cold Spring Harbor Symp. Quant. Biol.* **LXII**, 493-504.

**Attié, T., Till, M., Pelet, A., Amiel, J., Edery, P., Boutrand, L., Munnich, A. and Lyonnet, S.** (1995). Mutation of the endothelin-receptor B gene in Waardenburg-Hirschsprung disease. *Hum. Mol. Genet.* **4**, 2407-2409.

**Baynash, A. G., Hosoda, K., Giaid, A., Richardson, J. A., Emoto, N.,**

**Hammer, R. E. and Yanagisawa, M.** (1994). Interaction of endothelin-3 with endothelin-B receptor is essential for development of epidermal melanocytes and enteric neurons. *Cell* **79**, 1277-1285.

**Bolande, R. P.** (1974). The neurocristopathies: a unifying concept of disease arising in neural crest maldevelopment. *Hum. Pathol.* **5**, 409-429.

**Bondurand, N., Kobetz, A., Pingault, V., Lemort, N., EnchaRazavi, F., Couly, G., Goerich, D. E., Wegner, M., Abitbol, M. and Goossens, M.** (1998). Expression of the SOX10 gene during human development. *FEBS Lett.* **432**, 168-172.

**Bondurand, N., Pingault, V., Goerich, D. E., Lemort, N., Sock, E., Caignec, C. L., Wegner, M. and Goossens, M.** (2000). Interaction among SOX10, PAX3 and MITF, three genes altered in Waardenburg syndrome. *Hum. Mol. Genet.* **9**, 1907-1917.

**Britsch, S., Goerich, D. E., Riethmacher, D., Peirano, R. I., Rossner, M., Nave, K. A., Birchmeier, C. and Wegner, M.** (2001). The transcription factor Sox10 is a key regulator of peripheral glial development. *Genes Dev.* **15**, 66-78.

**Dutton, K., Dutton, J. R., Pauliny, A. and Kelsh, R. N.** (2001). A morpholino phenocopy of the *colourless* mutant. *Genesis* **30**, 188-189.

**Edery, P., Atti'e, T., Amiel, J., Pelet, A., Eng, C., Hofstra, R. M., Martelli, H., Bidaud, C., Munnich, A. and Lyonnet, S.** (1996). Mutation of the endothelin-3 gene in the Waardenburg-Hirschsprung disease (Shah-Waardenburg syndrome). *Nat. Genet.* **12**, 442-444.

**Haffter, P., Granato, M., Brand, M., Mullins, M., Hammerschmidt, M., Kane, D., Odenthal, J., van Eeden, F., Jiang, Y., Heisenberg, C. et al.** (1996). The identification of genes with unique and essential functions in the development of the zebrafish, Danio rerio. *Development* **123**, 1-36.

**Hailoran, M., Sato-Maeda, M., Warren, J., Su, F., Lele, Z., Krone, P., Kuwada, J. and Shoji, W.** (2000). Laser-induced gene expression in specific cells of transgenic zebrafish. *Development* **127**, 1953-1960.

**Herbarth, B., Pingault, V., Bondurand, N., Kuhlbrodt, K., Hermans-Borgmeyer, I., Puliti, A., Lemort, N., Goossens, M. and Wegner, M.** (1998). Mutation of the Sry-related Sox10 gene in Dominant megacolon, a mouse model for human Hirschsprung disease. *Proc. Natl. Acad. Sci. USA* **95**, 5161-5165.

**Hofstra, R. M., Osinga, J., Tan Sindhunata, G., Wu, Y., Kamsteeg, E. J., Stulp, R. P., van Ravenswaaij Arts, C., Majoor Krakauer, D., Angrist, M., Chakravarti, A. et al.** (1996). A homozygous mutation in the endothelin-3 gene associated with a combined Waardenburg type 2 and Hirschsprung phenotype (Shah-Waardenburg syndrome). *Nat. Genet.* **12**, 445-447.

**Hosoda, K., Hammer, R. E., Richardson, J. A., Baynash, A. G., Cheung, J. C., Giaid, A. and Yanagisawa, M.** (1994). Targeted and natural (piebald-lethal) mutations of endothelin-B receptor gene produce megacolon associated with spotted coat color in mice. *Cell* **79**, 1267-1276.

**Hukriede, N. A., Joly, L., Tsang, M., Miles, J., Tellis, P., Epstein, J. A., Barbazuk, W. B., Li, F. N., Paw, B., Postlethwait, J. H. et al.** (1999). Radiation hybrid mapping of the zebrafish genome. *Proc. Natl. Acad. Sci. USA* **96**, 9745-9750.

**Ikegami, R., Hunter, P. and Yager, T. D.** (1999). Developmental activation of the capability to undergo checkpoint-induced apoptosis in the early zebrafish embryo. *Dev. Biol.* **209**, 409-433.

**Kapur, R. P.** (1999). Early death of neural crest cells is responsible for total enteric aganglionosis in Sox10(Dom)/Sox10(Dom) mouse embryos. *Pediatr. Dev. Pathol.* **2**, 559-569.

**Kelsh, R. N., Brand, M., Jiang, Y.-J., Heisenberg, C. P., Lin, S., Haffter, P., Odenthal, J., Mullins, M. C., van Eeden, F. J., Furutani-Seiki, M. et al.** (1996). Zebrafish pigmentation mutations and the processes of neural crest development. *Development* **123**, 369-389.

**Kelsh, R. N. and Eisen, J. S.** (2000). The zebrafish *colourless* gene regulates development of non-ectomesenchymal neural crest derivatives. *Development* **127**, 515-525.

**Kelsh, R. N., Dutton, K., Medlin, J. and Eisen, J. S.** (2000a). Expression of zebrafish *fkd6* in neural crest-derived glia. *Mech. Dev.* **93**, 161-164.

**Kelsh, R. N., Schmid, B. and Eisen, J. S.** (2000b). Genetic analysis of melanophore development in zebrafish embryos. *Dev. Biol.* **225**, 277-293.

**Kimmel, C. B., Ballard, W. W., Kimmel, S. R., Ullmann, B. and Schilling, T. F.** (1995). Stages of embryonic development of the zebrafish. *Dev. Dyn.* **203**, 253-310.

**Knapik, E., Goodman, A., Atkinson, O., Roberts, C., Shiozawa, M., Sim, C., Weksler-Zangen, S., Trolliet, M., Futrell, C., Innes, B. et al.** (1996). A reference cross DNA panel for zebrafish (Danio rerio) anchored with simple sequence length polymorphisms. *Development* **123**, 451-460.

**Kuhlbrodt, K., Herbarth, B., Sock, E., Hermans-Borgmeyer, I. and Wegner, M.** (1998a). Sox10, a novel transcriptional modulator in glial cells. *J. Neurosci.* **18**, 237-250.

**Kuhlbrodt, K., Schmidt, C., Sock, E., Pingault, V., Bondurand, N., Goossens, M. and Wegner, M.** (1998b). Functional analysis of Sox10 mutations found in human Waardenburg-Hirschsprung patients. *J. Biol. Chem.* **273**, 23033-23038.

**Lane, P. W. and Liu, H. M.** (1984). Association of megacolon with a new dominant spotting gene (Dom) in the mouse. *J. Hered.* **75**, 435-439.

**Le Douarin, N. M.** (1982). *The Neural Crest*. Cambridge: Cambridge University Press.

**Lee, M., Goodall, J., Verastegui, C., Ballotti, R. and Goding, C. R.** (2000). Direct regulation of the microphthalmia promoter by Sox10 links Waardenburg-Shah syndrome (WS4)-associated hypopigmentation and deafness to WS2. *J. Biol. Chem.* **275**, 37978-37983.

**Lister, J. A., Robertson, C. P., Lepage, T., Johnson, S. L. and Raible, D. W.** (1999). nacre encodes a zebrafish microphthalmia-related protein that regulates neural-crest-derived pigment cell fate. *Development* **126**, 3757-3767.

**Liu, Q., Melnikova, I. N., Hu, M. and Gardner, P. D.** (1999). Cell type-specific activation of neuronal nicotinic acetylcholine receptor subunit genes by Sox10. *J. Neurosci.* **19**, 9747-9755.

**Malicki, J., Schier, A. F., Solnica-Krezel, L., Stemple, D. L., Neuhauss, S. C. F., Stainier, D. Y. R., Abdelilah, S., Rangini, Z., Zwartkruis, F. and Driever, W.** (1996). Mutations affecting development of the zebrafish ear. *Development* **123**, 275-283.

**Marusich, M. F., Furneaux, H. M., Henion, P. D. and Weston, J. A.** (1994). Hu neuronal proteins are expressed in proliferating neurogenic cells. *J. Neurobiol.* **25**, 143-155.

**Mayer, T. C.** (1965). The development of piebald spotting in mice. *Dev. Biol.* **11**, 319-334.

**Mayer, T. C. and Maltby, E. L.** (1964). An experimental investigation of pattern development in lethal spotting and belted mouse embryos. *Dev. Biol.* **9**, 269-286.

**Odenthal, J. and Nuesslein-Volhard, C.** (1998). *fork head* domain genes in zebrafish. *Dev. Genes Evol.* **208**, 245-258.

**Opdecamp, K., Nakayama, A., Nguyen, M. T. T., Hodgkinson, C. A., Pavan, W. J. and Arnheiter, H.** (1997). Melanocyte development in vivo and in neural crest cell cultures: Crucial dependence on the Mitf basic-helix-loop-helix-zipper transcription. *Development* **124**, 2377-2386.

**Parichy, D. M., Rawls, J. F., Pratt, S. J., Whitfield, T. T. and Johnson, S. L.** (1999). Zebrafish sparse corresponds to an orthologue of c-kit and is required for the morphogenesis of a subpopulation of melanocytes, but is not essential for hematopoiesis or primordial germ cell development. *Development* **126**, 3425-3436.

**Parichy, D. M., Ransom, D. G., Paw, B., Zon, L. I. and Johnson, S. L.** (2000). An orthologue of the kit-related gene fms is required for development of neural crest-derived xanthophores and a subpopulation of adult melanocytes in the zebrafish, Danio rerio. *Development* **127**, 3031-3044.

**Pattyn, A., Morin, X., Cremer, H., Goridis, C. and Brunet, J.-F.** (1999). The homeobox gene *Phox2b* is essential for the development of autonomic neural crest derivatives. *Nature* **399**, 366-370.

**Peirano, R. I. and Wegner, M.** (2000). The glial transcription factor Sox10 binds to DNA both as monomer and dimer with different functional consequences. *Nucleic Acids Res.* **28**, 3047-3055.

**Pevny, L. H. and Lovell-Badge, R.** (1997). Sox genes find their feet. *Curr. Opin. Genet Dev.* **7**, 338-44.

**Pingault, V., Bondurand, N., Kuhlbrodt, K., Goerich, D. E., Prehu, M. O., Puliti, A., Herbarth, B., Hermans-Borgmeyer, I., Legius, E., Matthijs, G. et al.** (1998). SOX10 mutations in patients with Waardenburg-Hirschsprung disease. *Nat. Genet.* **18**, 171-173.

**Potterf, S. B., Furumura, M., Dunn, K. J., Arnheiter, H. and Pavan, W. J.** (2000). Transcription factor hierarchy in Waardenburg syndrome: regulation of MITF expression by SOX10 and PAX3. *Hum. Genet.* **107**, 1-6.

**Puffenberger, E. G., Hosoda, K., Washington, S. S., Nakao, K., deWit, D., Yanagisawa, M. and Chakravarti, A.** (1994). A missense mutation of the endothelin-B receptor gene in multigenic Hirschsprung's disease. *Cell* **79**, 1257-1266.

**Pusch, C., Hustert, E., Pfeifer, D., Sudbeck, P., Kist, R., Roe, B., Wang, Z., Balling, R., Blin, N. and Scherer, G.** (1998). The SOX10/Sox10 gene from human and mouse: sequence, expression, and transactivation by the encoded HMG domain transcription factor. *Hum. Genet.* **103**, 115-123.

**Raible, D. W. and Eisen, J. S.** (1994). Restriction of neural crest cell fate in the trunk of the embryonic zebrafish. *Development* **120**, 495-503.

**Raible, D. W., Wood, A., Hodsdon, W., Henion, P. D., Weston, J. A. and Eisen, J. S.** (1992). Segregation and early dispersal of neural crest cells in the embryonic zebrafish. *Dev. Dyn.* **195**, 29-42.

**Rauch, G. J., Hammerschmidt, M., Blader, P., Schauerte, H. E., Strahle, U., Ingham, P. W., McMahon, A. P. and Haffter, P.** (1997). Wnt5 is required for tail formation in the zebrafish embryo. *Cold Spring Harbor Symp. Quant. Biol.* **62**, 227-234.

**Reyes, R.** (1999). Development and Death of Zebrafish Rohon-Beard Spinal Sensory Neurons. In *Institute of Neuroscience*, pp. 1-99. Eugene, OR: University of Oregon.

**Schilling, T. F. and Kimmel, C. B.** (1994). Segment and cell type lineage restrictions during pharyngeal arch development in the zebrafish embryo. *Development* **120**, 483-494.

**Schulte-Merker, S., Van Eeden, F. J. M., Halpern, M. E., Kimmel, C. B. and Nüsslein-Volhard, C.** (1994). *no tail* (*ntl*) is the zebrafish homologue of the mouse *T* (*Brachyury*) gene. *Development* **120**, 1009-1015.

**Smith, M., Hickman, A., Amanze, D., Lumsden, A. and Thorogood, P.** (1994). Trunk neural crest origin of caudal fin mesenchyme in the zebrafish *Brachydanio rerio*. *Proc. R. Soc. London B Biol. Sci.* **256**, 137-145.

**Southard-Smith, E. M., Kos, L. and Pavan, W. J.** (1998). Sox10 mutation disrupts neural crest development in DOM Hirschsprung mouse model. *Nat. Genet.* **18**, 60-64.

**Southard-Smith, E. M., Angrist, M., Ellison, J. S., Agarwala, R., Baxevanis, A. D., Chakravarti, A. and Pavan, W. J.** (1999). The Sox10(Dom) mouse: modeling the genetic variation of Waardenburg-Shah (WS4) syndrome. *Genome Res.* **9**, 215-225.

**Strimmer, K. and von Haeseler, A.** (1996). Quartet puzzling: A quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**, 964-969.

**Tachibana, M., Takeda, K., Nobukuni, Y., Urabe, K., Long, J. E., Meyers, K. A., Aaronson, S. A. and Miki, T.** (1996). Ectopic expression of MITF, a gene for Waardenburg syndrome type 2, converts fibroblasts to cells with melanocyte characteristics. *Nature Genet.* **14**, 50-54.

**Tamura, K. and Nei, M.** (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial-DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**, 512-526.

**Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. and Higgins, D. G.** (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **25**, 4876-4882.

**Verastegui, C., Bille, K., Ortonne, J. P. and Ballotti, R.** (2000). Regulation of microphthalmia-associated transcription factor gene by the waardenburg syndrome type 4 Gene, Sox10. *J. Biol. Chem.* **275**, 30757-60.

**Wegner, M.** (1999). From head to toes: the multiple facets of Sox proteins. *Nucleic Acids Res.* **27**, 1409-1420.

**Wehrle-Haller, B. and Weston, J. A.** (1995). Soluble and cell-bound forms of steel factor activity play distinct roles in melanocyte precursor dispersal and survival on the lateral neural crest migration pathway. *Development* **121**, 731-742.

**Wehrle-Haller, B. and Weston, J. A.** (1997). Receptor tyrosine kinase-dependent neural crest migration in response to differentially localized growth factors. *BioEssays* **19**, 337-345.

**Yuan, H., Corbi, N., Basilico, C. and Dailey, L.** (1995). Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 and Oct-3. *Genes Dev.* **9**, 2635-2645.

genesis 30:188-189 (2001)

# A Morpholino Phenocopy of the *colourless* Mutant

**K. Dutton,* J. R. Dutton, A. Pauliny, and R. N. Kelsh**

Department of Biology and Biochemistry, University of Bath, Bath, United Kingdom

Received 12 April 2001; Accepted 29 May 2001
**Published online 23 July 2001; DOI 10.1002/gene.1062**

**Key words:** neural crest, *sox10*, Waardenburg-Shah syndrome, pigmentation, neurocristopathy

We have utilized the modified antisense RNA morpholino technology to effectively phenocopy zebrafish *colourless/sox10* (*cls*) mutations. The *cls* locus was identified in mutagenesis screens (Kelsh *et al.*, 1996; Malicki *et al.*, 1996). Homozygous mutants are characterized by extensive defects in nonectomesenchymal fates (neurons, glia, and pigment cells) derived from the neural crest (Kelsh and Eisen 2000; Kelsh *et al.*, 2000a, 2000b). Additionally, *cls* mutants have small otic vesicles (Whitfield *et al.*, 1996). Zebrafish *cls* mutants are models for two human neurocristopathies, Hirschsprung's disease, characterized by few or no enteric ganglia, and Waardenburg-Shah syndrome, which combines Hirschsprung's disease with pigment defects. We have recently shown that the *cls* mutant phenotype results from disruptions in the zebrafish *sox10* homologue and that *sox10* expression is first seen at approximately 11 h postfertilization (hpf) in premigratory neural crest cells (A. Pauliny and R. N. Kelsh, unpublished). The *cls* mutant phenotype can first be detected at 21 hpf in dopachrome tautomerase (dct) in situ or scored visibly at 27 hpf (Kelsh *et al.*, 2000a). Thus, we demonstrate that morpholino oligos can effectively phenocopy late embryonic phenotypes in zebrafish. We have characterized morphants generated using morpholino oligos designed to target the zebrafish *sox10* homologue. The phenotypes resulting from injection of these morpholinos are consistent with the *cls* mutant phenotype affecting both neural crest derivatives and otic vesicles with an optimal response achieved with a dose of 16.5 ng (Fig 1 and Table 1). In strong mutant phenotypes, no normal melanophores are seen. However, even with high doses of morpholino, some normal melanophores are present, implying that the characterized strong mutant alleles are likely nulls. At lower doses, all morphant embryos are reminiscent of the single reported weak *cls* allele, which is not presently available (Malicki *et al.*, 1996). As most known *cls* alleles display strong phenotypes, the ability to produce a graded series of *cls* hypomorphs with morpholinos will be invaluable in examining the role of *sox10* in neural crest development.

**Table 1**
*Dose–Response Curve*

| Dose | Weak phenocopy[a] | Strong phenocopy[b] | *n* |
|---|---|---|---|
| 3 ng | 10% | — | 19 |
| 4.5 ng | 12% | — | 105 |
| 6 ng | 21% | — | 105 |
| 7 ng | 26% | — | 98 |
| 9 ng | 18% | 11% | 104 |
| 14 ng | 14% | 30% | 59 |
| 16.5 ng | 44% | 24% | 218 |
| 37 ng | 49% | 29% | 43 |

[a]Weak phenocopies had at least 25 melanophores with a mean of 67 (SD ± 25).

[b]Strong phenocopies had a mean of 11 (SD ± 8) but less than 25 total melanophores. One to eight cell AB wild-type embryos were injected with 4.6 nl of morpholino oligo diluted as recommended by Gene Tools, LLC and incubated at 28.5°C. Embryos with more than 50% of wild-type pigment at 48 hpf were scored as wild type due to variations in embryonic development. At doses over 16.5 ng, nonspecific deformities became apparent.

## LITERATURE CITED

Kelsh RN, Eisen JS. 2000. The zebrafish *colourless* gene regulates development of non-ectomesenchymal neural crest derivatives. Development 127:515-525.

Kelsh RN, Brand M, Jiang Y-J, Heisenberg CP, Lin S, Haffter P, Odenthal J, Mullins MC, van Eeden FJ, Furutani-Seiki M, Granato M, Hammerschmidt M, Kane DA, Warga RM, Beuchle D, Vogelsang L, Nusslein-Volhard C. 1996. Zebrafish pigmentation mutations and the processes of neural crest development. Development 123: 369-389.

Kelsh RN, Dutton K, Medlin J, Eisen JS. 2000a. Expression of zebrafish *fkd6* in neural crest-derived glia. Mech Dev 93:161-164.

Kelsh RN, Schmid B, Eisen JS. 2000b. Genetic analysis of melanophore development in zebrafish embryos. Dev Biol 225:277-293.

Malicki J, Schier AF, Solnica-Krezel L, Stemple DL, Neuhauss SCF, Stainier DYR, Abdelilah S, Rangini Z, Zwartkruis F, Driever W. 1996. Mutations affecting development of the zebrafish ear. Development 123:275-283.

Marusich MF, Furneaux HM, Henion PD, Weston JA. 1994. Hu neuronal proteins are expressed in proliferating neurogenic cells. J Neurobiol 25:143-155.

Whitfield TT, Granato M, van Eeden FJ, Schach U, Brand M, Furutani-Seiki M, Haffter P, Hammerschmidt M, Heisenberg CP, Jiang YJ, Kane DA, Kelsh RN, Mullins MC, Odenthal J, Nusslein-Volhard C. 1996. Mutations affecting development of the zebrafish inner ear and lateral line. Development 123:241-254.

* Correspondence to: R. N. Kelsh, Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK.
E-mail: bssrnk@bath.ac.uk

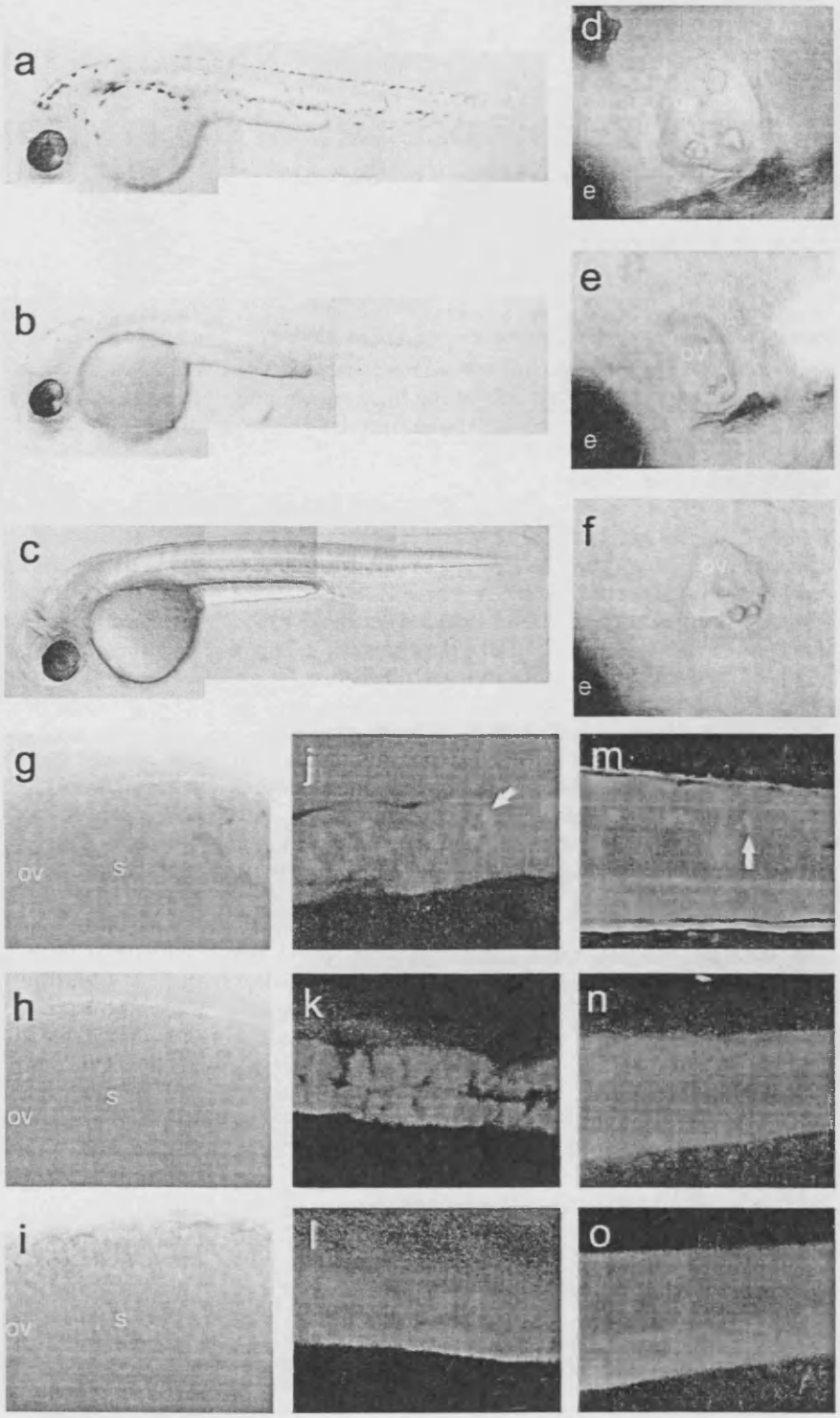



**FIG. 1.** *sox10* knockdown embryos phenocopy *cls* mutant phenotypes. (**a**, **d**, **g**, **j**, **m**) Uninjected wild-type embryos shown in comparison with embryos injected with 16.5 ng of a morpholino oligo (5′-GCCACAGGTGACTTCGGTAGGTTTA-3′) designed to target the −43 to −19 region of the (**b**, **e**, **h**, **k**, **n**) *sox10* sequence and (**c**, **f**, **i**, **l**, **o**) *cls* mutant embryos. A morpholino oligo (5′-ATGCTGTGCTCCTCCGCCGACATCG-3′) designed to target the −23 to +2 region of the *sox10* sequence gave similar phenocopies (data not shown). Panels show lateral views of (**a**–**f**) live whole-mounts, (**g**–**i**) fixed embryos processed for *dct* (Kelsh *et al.*, 2000b) in situ hybridization, or (**j**–**o**) anti-Hu mAb 16A11 (Marusich *et al.*, 1994) antibody staining. (**d**–**f**) Lateral views of the otic vesicle of a (**d**) 72 hpf wild-type embryo with the reduced otic vesicle, (**e**) small otoliths of an injected sibling, and (**f**) *cls* mutant. The *dct* in situ hybridization reveals melanoblasts in the anterior trunk of a (**g**) 22 hpf wild-type embryo; these are absent in an (**h**) injected sibling, and (**i**) *cls* mutant. Anti-Hu staining demonstrates a significant decrease in the number of enteric neurons (arrows) in the hindgut of an (**k**) injected 5-day postfertilization (dpf) embryo compared with a (**j**) wild-type sibling and similar to a (**l**) *cls* mutant. Hu-positive dorsal root ganglia (arrow) in the tail of a (**m**) 5 dpf wild-type embryo are lacking in an (**n**) injected sibling, and (**o**) *cls* mutant (e, eye; ov, otic vesicle; s, somite).